

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OW protein - protein search, using sw model

Run on: September 30, 2004, 10:58:43 / Search time 16.5905 Seconds
(without alignments)
150.748 Million cell updates/sec

Title: US-09-988-851A-12

Perfect score: 135
Sequence: 1 TWLEERDEWVRSIAVDQAQHAARRVAS 26

Scoring table: BLOSUM62
Gapop 10.0, Gapext 0.5

Searched: 283366 seqs, 96191526 residues

Total number of hits satisfying chosen parameters: 283366

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database: PIR 78:*

1: PIR1: *
2: PIR2: *
3: PIR3: *
4: PIR4: *

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	132	97.8	798	2	S09867
2	54	40.0	129	2	AF0531
3	53.5	39.6	805	2	T49385
4	51	37.8	2470	2	S57085
5	51	37.8	2473	1	S38040
6	50.5	37.4	548	2	JCS030
7	50.5	37.4	549	2	G83345
8	50.5	37.4	861	2	H83012
9	50	37.0	203	2	H82653
10	50	37.0	331	2	F83346
11	50	37.0	384	2	H91130
12	50	37.0	384	2	H85975
13	50	37.0	451	2	S30401
14	49	36.3	347	2	C87323
15	49	36.3	608	2	P83357
16	49	36.3	940	2	D87006
17	49	36.3	2388	2	JB0271
18	48.5	35.9	903	2	T26743
19	48.5	35.9	951	2	T26737
20	48.5	35.9	975	2	T26737
21	48.5	35.9	1758	2	T34393
22	48	35.6	719	2	T12258
23	48	35.6	1404	2	E30658
24	48	35.6	1404	2	E85509
25	47.5	35.2	609	2	T45497
26	47	34.8	151	2	D86891
27	47	34.8	232	2	A83208
28	47	34.8	350	1	S00755
29	47	34.8	368	2	I50233

30	47	34.8	448	2	I48128	vimentin - Chinese
31	47	34.8	460	2	A29329	vimentin - chicken
32	47	34.8	464	1	VEHY	vimentin - golden
33	47	34.8	466	2	A25074	vimentin - human
34	47	34.8	466	2	A43803	vimentin - mouse
35	47	34.8	466	2	S22119	vimentin - rat
36	47	34.8	666	2	F71310	probable periplasm
37	47	34.8	1524	2	T30518	cardamyl-phosphat
38	47	34.8	1870	2	D88486	protein F20H11.2 l
39	47	34.8	2513	2	G96536	hypothetical prote
40	46.5	34.4	99	2	AC3755	conserved hypothet
41	46.5	34.4	104	2	F97536	hypothetical prote
42	46.5	34.4	660	2	F72610	probable acetyl-co
43	46	34.1	299	2	T21685	hypothetical prote
44	46	34.1	430	1	VEMSGF	glial fibrillary a
45	46	34.1	450	2	I50484	vimentin beta - go

ALIGNMENTS

RESULT 1

S09867 hypothetical protein U1.02 - human cytomegalovirus (strain AD169)

C/Species: human cytomegalovirus, human herpesvirus 5

C/Date: 07-Sep-1990 #sequence_revision 07-Sep-1990 #text_change 08-Oct-1999

C/Accession: S09867

R/Chen, M.S.; Bankier, A.T.; Beck, S.; Bohm, R.; Brown, C.M.; Cerny, R.; Horsnell, T.; J

M.; Barré, B.G.

Curr. Top. Microbiol. Immunol. 154, 125-169, 1990

A/Title: Analysis of the protein-coding content of the sequence of human cytomegalovirus

A/Reference number: S09749; NCID:50269039; PMID:216119

A/Accession: S09867

A/Status: nucleic acid sequence not shown; translation not shown

A/Molecule type: DNA

A/Residues: 1-798 <CHE>

A/Cross-references: EMBL:X17403; NID:959591; PID:CA5338.1; PID:e27303; PID:g1780881

A/Note: this sequence was submitted to the EMBL Data Library, December 1989

A/Note: this reading frame extends between two stop codons and does not begin with a start

A/Note: this sequence was submitted to the EMBL Data Library, December 1989

A/Accession: AF0531

R/Perkhill, J.; Dougan, G.; James, K.D.; Thomson, N.R.; Pickard, D.; Main, J.; Churcher, T.;

Th, T.; Connerton, P.; Cronin, A.; Davis, P.; Davies, R.M.; Dowd, L.; White, N.; Farrar, S.;

Nature 413, 848-852, 2001

A/Authors: Parry, C.; Quail, M.; Rutherford, K.; Simmonds, M.; Skelton, J.; Stevens, K.;

A/Title: Complete genome sequence of a multiple drug resistant *Salmonella enterica* serov

A/Reference number: AB0502; NCID:21534947; PMID:11677608

A/Accession: AF0531

A/Status: preliminary

A/Molecule type: DNA

A/Residues: 1-120 <PAR>

A/Cross-references: GB:AL513382; PID:CAD08694.1; PID:g16501517; GSPDB:GN00176

A/Genes: STY0260

Query Match 40.0%; Score 54; DB 2; Length 129;

Best Local Similarity 50.0%; Pred. No. 1.4;
Matches 9; Conservative 3; Mismatches 6; Indels 0; Gaps 0;

Qy 2 TWLEERDEWVSLAVDAQH 19
Db 37 WREERDSWKQDLALNGQV 54

RESULT 3

149385
hypothetical protein B1D1.390 [imported] - Neurospora crassa
C:Species: Neurospora crassa
C:Date: 02-Jun-2000 #sequence_revision 02-Jun-2000 #text_change 02-Jun-2000
C:Accession: T49385
R:Schulze, U.; Aign, V.; Honeisel, J.; Brandt, P.; Fattmann, B.; Holland, R.; Nyakatura, submitted to the Protein Sequence Database, May 2000
A:Reference number: 225022
A:Accession: T49385
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-805 <SCCH>
A:Cross-references: EMBL:ALJ55927; GSPDB:GN00116; NCSP:B1D1.390
A:Experimental source: BAC clone B1D1; strain OR74A
C:Genetics:
A:Gene: NCSP:B1D1.390
A:Map position: 6
A:introns: 54/3; 212/3

Query Match 39.6%; Score 53.5; DB 2; Length 805;
Best Local Similarity 43.5%; Pred. No. 11;
Matches 10; Conservative 5; Mismatches 7; Indels 1; Gaps 1;

Qy 1 TWLEERDEWVSL-AYDAQHAR 22
Db 601 SMWVEQSDWVSLAIDSDGHR 623

RESULT 4

557085
1-phosphatidylinositol 3-kinase (EC 2.7.1.137) TOR1 - yeast (Saccharomyces cerevisiae)
N:Alternate names: protein J1803; protein YJR066w
C:Species: Saccharomyces cerevisiae
C:Date: 13-Sep-1995 #sequence_revision 01-Mar-1996 #text_change 18-Jun-1999
C:Accession: S57085; A54428; S43940; S71688
R:Manus, V.; Huang, M.E.; Galibert, F.
A:Reference number: S57085
A:Reference number: S57085
A:Molecule type: DNA
A:Residues: 1-2470 <MAN>
A:Cross-references: EMBL:Z49566; NID:g1015742; PIDN:CAA89594.1; PID:g1015743; MIPS:YJR06
R:Caletkey, R.; Young, P.R.; McLaughlin, M.M.; Bergsma, D.D.; Koltin, Y.; Sathe, G.M.; Mol. Cell. Biol. 13, 6012-6023, 1993
A:Title: Dominant missense mutations in a novel yeast protein related to mammalian phosphatidylinositol 3-kinase
A:Reference number: A54428; MUID:94019276; PMID:8413204
A:Accession: A54428
A:Status: nucleic acid sequence not shown

A:Molecule type: DNA
A:Residues: 1-57, 'G', 59-114, 'T', 116-132, 'N', 134-395, 'K', 397-546, 'S', 548-1467, 'R', 1469-16
A:Cross-references: GB:119540; NID:g408955; PIDN:AAB66881.1; PID:g408956
A:Note: the authors translated the codon CCG for residue 1468 as Ala
R:Halliwell, S.B.; Wagner, P.; Kunz, J.; Deuter-Reinhard, M.; Henriquez, R.; Hall, M.N. Mol. Biol. Cell 5, 105-118, 1994
A:Title: TOR1 and TOR2 are structurally similar but not identical phosphatidylinositol 3-kinases
A:Reference number: S43940; MUID:94243030; PMID:8186460
A:Accession: S43940
A:Molecule type: DNA
A:Residues: 1-230, 'R', 233-395, 'K', 397-546, 'S', 548-674, 'T', 676-1291, 'E', 1293-1435, 'A', 1435-2470 <HEHL>
A:Cross-references: EMBL:X74857; NID:g468738; PIDN:CAA52849.1; PID:g468739
R:Huang, M.E.; Manus, V.; Chuet, J.C.; Galibert, F.
A:Title: Analysis of a 62 kb DNA sequence of chromosome X reveals 36 open reading frames

A:Reference number: S71676; MUID:96437976; PMID:8840504

A:Accession: S71688

A:Status: nucleic acid sequence not shown; translation not shown

A:Molecule type: DNA

A:Residues: 1-2470 <HUA>
A:Cross-references: EMBL:L47993; NID:g1019675; PIDN:AAB39292.1; PID:g1019668
A:Note: the nucleotide sequence was submitted to the EMBL Data Library, October 1995

C:Genetics:
A:Gene: SGD:TOR1; DR1

A:Cross-references: SGD:S0003827; MIPS:YJR066w

A:Map position: 10R

A:Function:

A:Description: required for translation initiation; required for G1 progression
C:Superfamily: Yeast TOR2 protein

C:Keywords: DNA binding; EF hand; leucine zipper; phosphotransferase

Query Match 37.8%; Score 51; DB 2; Length 2470;
Best Local Similarity 66.7%; Pred. No. 84;
Matches 10; Conservative 1; Mismatches 4; Indels 0; Gaps 0;

Qy 1 TWLEERDEWVSLAV 15
Db 2249 TWLEERTTYTRSLAV 2263

RESULT 5

538040
1-phosphatidylinositol 3-kinase (EC 2.7.1.137) TOR2 - yeast (Saccharomyces cerevisiae)
N:Alternate names: protein YKL203C
C:Species: Saccharomyces cerevisiae
C:Date: 10-Sep-1999 #sequence_revision 10-Sep-1999 #text_change 12-Nov-1999
C:Accession: S38040; S38041; S31106
R:Maia e Silva, A.; Bossier, P.; Valleja, C.; Fernandes, L.; Soares, H.; Guerreiro, P.; R submitted to the Protein Sequence Database, March 1994
A:Reference number: S38024
A:Accession: S38040
A:Molecule type: DNA
A:Residues: 1-2360 <MAI>
A:Cross-references: EMBL:Z28203; GSPDB:GN00011; MIPS:YKL203C
A:Experimental source: strain S288C
R:Pohl, T.M.; Pohl, P.M.
A:Reference number: S37897
A:Reference number: S37897
A:Accession: S38041
A:Molecule type: DNA
A:Residues: 1902-2473 <POH>
A:Cross-references: EMBL:Z28203; GSPDB:GN00011; MIPS:YKL203C
A:Experimental source: strain S288C
R:Kunz, J.; Henriquez, R.; Schneider, U.; Deuter-Reinhard, M.; Movva, N.R.; Hall, M.N. Cell 73, 585-596, 1993
A:Title: Target of rapamycin in yeast, TOR2, is an essential phosphatidylinositol kinase
A:Reference number: S35106; MUID:93258821; PMID:8387896
A:Accession: S35106
A:Molecule type: DNA
A:Residues: 1-1472, 'G', 1473-2473 <KUN>
A:Cross-references: EMBL:X71416; NID:g298027; PIDN:CAA50548.1; PID:g298028
C:Genetics:
A:Gene: SGD:TOR2; MIPS:YKL203C
A:Cross-references: SGD:S0001686; MIPS:YKL203C
A:Map position: 11L
C:Superfamily: Yeast TOR2 protein
C:Keywords: phosphotransferase

Query Match 37.8%; Score 51; DB 1; Length 2473;
Best Local Similarity 66.7%; Pred. No. 84;
Matches 10; Conservative 1; Mismatches 4; Indels 0; Gaps 0;

Qy 1 TWLEERDEWVSLAV 15
Db 2252 TWLEERTTYTRSLAV 2266

RESULT 6

JC5090
pyoverdine synthetase component E - Pseudomonas aeruginosa
C:Accession: H83012
C:Species: Pseudomonas aeruginosa
C>Date: 31-Jan-1997 #sequence_revision 31-Jan-1997 #text_change 02-Feb-2001
C:Accession: JC5090; S53997; S54001
R:Stover, C.K.; Pham, X.Q.; Erwin, A.L.; Mizoguchi, S.D.; Warren, P.; Hickey, M.J.; Br
adman, S.; Yuan, Y.; Brody, L.L.; Coulter, S.N.; Folger, K.R.; Kas, A.; Laidig, K.; Lim,
J.; Lory, S.; Olson, M.V.
A:Title: Characterization of the pvdE gene which is required for pyoverdine synthesis in
Pseudomonas aeruginosa
A:Reference number: JC5090; MUID:57075905; PMID:8918232
A:Accession: JC5090
A:Molecule type: DNA
A:Residues: 1-548 <MCM>
A:Cross-references: GB:U07359; NID:g1633044; PIDN:AA60200.1; PID:g466460
R:Rombel, I.T.; McMorran, B.J.; Lamont, I.L.
Mol. Gen. Genet. 246: 519-528, 1995
A:Title: Identification of a DNA sequence motif required for expression of iron-regulate
d pyoverdine synthetase component E in Pseudomonas aeruginosa
A:Reference number: S53996; MUID:95198690; PMID:7891666
A:Accession: S53997
A:Molecule type: DNA
A:Residues: 1-80 <ROM>
A:Cross-references: EMBL:U07359
C:Comment: This enzyme is a member of the ATP-binding cassette family of membrane transp
orters
C:Genetics:
A:Gene: pvdE
C:Superfamily: unassigned ATP-binding cassette proteins; ATP-binding cassette homology
C:Keywords: ATP; nucleotide binding; P-loop
C:35-539/Domain: ATP-binding cassette homology <ABC>
F:372-379/Region: nucleotide-binding motif A (P-loop)

Query Match 37.4%; Score 50.5; DB 2; Length 548;
Best Local Similarity 50.0%; Pred. No. 20;
Matches 14; Conservative 0; Mismatches 5; Indels 9; Gaps 2;

OY 2 WLEBRSLAVDAQHARRY 24
DB 478 WLEBRPVLVDEW----AADODPAFRRY 501

RESULT 7
G83345
pyoverdine biosynthesis protein PvdE PA2397 [imported] - Pseudomonas aeruginosa (strain
C:Species: Pseudomonas aeruginosa
C>Date: 15-Sep-2000 #sequence_revision 15-Sep-2000 #text_change 17-May-2002
C:Accession: G83345
R:Stover, C.K.; Pham, X.Q.; Erwin, A.L.; Mizoguchi, S.D.; Warren, P.; Hickey, M.J.; Br
adman, S.; Yuan, Y.; Brody, L.L.; Coulter, S.N.; Folger, K.R.; Kas, A.; Laidig, K.; Lim,
J.; Lory, S.; Olson, M.V.
A:Title: Complete genome sequence of Pseudomonas aeruginosa PA01, an opportunistic patho
gen
A:Reference number: A82950; MUID:20437337; PMID:10984043
A:Accession: G83345
A:Molecule type: DNA
A:Status: preliminary
A:Residues: 1-549 <STO>
A:Cross-references: GB:AE004666; GB:AE004091; NID:g948438; PIDN:AA605785.1; GSPDB:GN001
C:Experimental source: strain PA01
A:Genetics:
A:Gene: pvdE; PA2397
C:Superfamily: unassigned ATP-binding cassette proteins; ATP-binding cassette homology

Query Match 37.4%; Score 50.5; DB 2; Length 549;
Best Local Similarity 50.0%; Pred. No. 20;
Matches 14; Conservative 0; Mismatches 5; Indels 9; Gaps 2;

OY 2 WLEBRSLAVDAQHARRY 24
DB 479 WLEBRPVLVDEW----AADODPAFRRY 502

RESULT 8
H83012
periplasmic glucanase biosynthesis protein Mdh PA5077 [imported] - Pseudomonas aeruginosa
C:Species: Pseudomonas aeruginosa

C>Date: 15-Sep-2000 #sequence_revision 15-Sep-2000 #text_change 31-Dec-2000
C:Accession: H83012
R:Stover, C.K.; Pham, X.Q.; Erwin, A.L.; Mizoguchi, S.D.; Warren, P.; Hickey, M.J.; Br
adman, S.; Yuan, Y.; Brody, L.L.; Coulter, S.N.; Folger, K.R.; Kas, A.; Laidig, K.; Lim,
J.; Lory, S.; Olson, M.V.
A:Title: Complete genome sequence of Pseudomonas aeruginosa PA01, an opportunistic patho
gen
A:Reference number: A82950; MUID:20437337; PMID:10984043
A:Accession: H83012
A:Molecule type: DNA
A:Status: preliminary
A:Residues: 1-861 <STO>
A:Cross-references: GB:AE004920; GB:AE004091; NID:g9951358; PIDN:AA608462.1; GSPDB:GN001
C:Experimental source: strain PA01
A:Genetics:
A:Gene: mdh; PA5077

Query Match 37.4%; Score 50.5; DB 2; Length 861;
Best Local Similarity 43.5%; Pred. No. 32;
Matches 10; Conservative 3; Mismatches 7; Indels 3; Gaps 1;

OY 2 WLEBRDEWY---RSLAVDAQHARRY 21
DB 819 WLEBRDDWLGWRKXADDPHAR 841

RESULT 9
H82653
phage-related lysozyme Xf1669 Xf1564 [imported] - Xylella fastidiosa (strain 9a5c)
C:Species: Xylella fastidiosa
C>Date: 18-Aug-2000 #sequence_revision 20-Aug-2000 #text_change 02-Sep-2000
C:Accession: H82653; H82663
R:Anonymous, The Xylella fastidiosa Consortium of the Organization for Nucleotide Sequenc
ing
A:Title: The genome sequence of the plant pathogen Xylella fastidiosa.
A:Reference number: A82515; MUID:20365717; PMID:10910347
A:Note: for a complete list of authors see reference number A59328 below
A:Accession: H82653
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-203 <SIM>
A:Cross-references: GB:AE003992; GB:AE003849; NID:g9106715; PIDN:AA64478.1; GSPDB:GN001
A:Experimental source: strain 9a5c
A:Accession: H82663
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-203 <S12>
A:Cross-references: GB:AE003986; GB:AE003849; NID:g9106606; PIDN:AA64373.1; GSPDB:GN001
A:Experimental source: strain 9a5c
R:Stover, C.K.; Pham, X.Q.; Erwin, A.L.; Mizoguchi, S.D.; Warren, P.; Hickey, M.J.; Br
adman, S.; Yuan, Y.; Brody, L.L.; Coulter, S.N.; Folger, K.R.; Kas, A.; Laidig, K.; Lim,
J.; Lory, S.; Olson, M.V.
A:Title: Complete genome sequence of Pseudomonas aeruginosa PA01, an opportunistic patho
gen
A:Reference number: A82950; MUID:20437337; PMID:10984043
A:Accession: H83012
A:Molecule type: DNA
A:Status: preliminary
A:Residues: 1-549 <STO>
A:Cross-references: GB:AE004666; GB:AE004091; NID:g948438; PIDN:AA605785.1; GSPDB:GN001
C:Experimental source: strain PA01
A:Genetics:
A:Gene: pvdE; PA2397
C:Superfamily: unassigned ATP-binding cassette proteins; ATP-binding cassette homology

Query Match 37.0%; Score 50; DB 2; Length 203;
Best Local Similarity 48.0%; Pred. No. 8.2;
Matches 12; Conservative 4; Mismatches 7; Indels 2; Gaps 1;

OY 2 WLEBRDEWYSLAVDAQHARRYAS 26

Db 179 WUEFGQDWOAALA--AEHAAVKRAS 201

RESULT 10

F83346

Probable transmembrane sensor PA2388 [imported] - *Pseudomonas aeruginosa* (strain PAO1)
C:Species: *Pseudomonas aeruginosa*
C>Date: 15-Sep-2000 #sequence_revision 15-Sep-2000 #text_change 28-Jul-2003

C/Accession: F83346

R:Stover, C.K.; Pham, X.Q.; Erwin, A.L.; Micozuchi, S.D.; Warren, P.; Hickey, M.J.; Br
adman, S.; Yuan, Y.; Brody, L.L.; Coulter, S.N.; Folger, K.R.; Kas, A.; Lapid, K.; Lam,
.; Lory, S.; Olson, M.V.
Nature 406, 959-964, 2000

A:Title: Complete genome sequence of *Pseudomonas aeruginosa* PAO1, an opportunistic patho
A:Reference number: A82950, MUID:20437337, PMID:10984643

A/Accession: F83346

A/Status: preliminary

A/Molecule type: DNA

A/Residues: 1-331 <STO>

A/Cross-references: GB:AE004665; GB:AE004091, NID:9948426; PIDN:AA05776.1; GSPDB:GN001

A/Experimental source: strain PAO1

C/Genetics:

C/Superfamily: Fe2+-dicitrate sensor, transmembrane component

Query Match 37.0%; Score 50; DB 2; Length 331;
Best Local Similarity 61.1%; Pred. No. 14;

Matches 11; Conservative 2; Mismatches 5; Indels 0; Gaps 0;

Qy 4 EERDEWVRSIAVDAQHAA 21

Db 31 EERGEFORMLAADPRHAA 48

RESULT 11

H91130

hypothetical protein EC64016 [imported] - *Escherichia coli* (strain O157:H7, substrain R1
C/Species: *Escherichia coli*
C/Date: 18-Jul-2001 #sequence_revision 18-Jul-2001 #text_change 18-Jul-2001

C/Accession: H91130

R:Havashi, T.; Makino, K.; Ohnishi, M.; Kurokawa, K.; Ishii, K.; Yokoyama, K.; Han, C.G.
gasawara, N.; Yasunaga, T.; Kihara, S.; Shiba, T.; Hattori, M.; Shinagawa, H.

DNA Res. 8, 11-22, 2001

A:Title: Complete genome sequence of enterohemorrhagic *Escherichia coli* O157:H7 and geno
A:Reference number: A96629, MUID:21156231, PMID:11258796

A/Accession: H91130

A/Status: preliminary

A/Molecule type: DNA

A/Residues: 1-384 <HAV>

A/Cross-references: GB:BA000007, PIDN:BA037439.1, PID:G13363469; GSPDB:GN00154

A/Experimental source: strain O157:H7, substrain R1MD 0509952

C/Genetics:

A/Gene: EC64016

Query Match 37.0%; Score 50; DB 2; Length 384;
Best Local Similarity 44.4%; Pred. No. 16;

Matches 12; Conservative 3; Mismatches 6; Indels 6; Gaps 2;

Qy 1 TWLEER-----DEWVRSIA-VDAQHAA 21

Db 14 TWLEERHQPRAMIRSLTNIDALHSA 40

RESULT 12

H85975

hypothetical protein agas [imported] - *Escherichia coli* (strain O157:H7, substrain EDL93
C/Species: *Escherichia coli*
C/Date: 16-Feb-2001 #sequence_revision 16-Feb-2001 #text_change 14-Sep-2001

C/Accession: H85975

R:Perna, N.T.; Plunkett III, G.; Burland, V.; Mau, B.; Glasner, J.D.; Rose, D.J.; Mayhew
iller, L.; Groetbeck, E.J.; Davis, N.W.; Lim, A.; Dimalanta, E.; Potamousis, K.; Apodaca,
Nature 409, 529-533, 2001

A:Title: Genome sequence of enterohemorrhagic *Escherichia coli* O157:H7.

A/Reference number: A85480, MUID:21074935, PMID:11206551

A/Accession: H85975

A/Status: preliminary

A/Molecule type: DNA

A/Residues: 1-384 <STO>

A/Cross-references: GB:AE005174, NID:G12517724; PIDN:AA058268.1; GSPDB:GN00145; UMG:2445

A/Experimental source: strain O157:H7, substrain EDL933

C/Genetics:

A/Gene: agas

Query Match 37.0%; Score 50; DB 2; Length 384;
Best Local Similarity 44.4%; Pred. No. 16;

Matches 12; Conservative 3; Mismatches 6; Indels 6; Gaps 2;

Qy 1 TWLEER-----DEWVRSIA-VDAQHAA 21

Db 14 TWLEERHQPRAMIRSLTNIDALHSA 40

RESULT 13

S30401

hypothetical protein 2 - *Streptomyces clavuligerus* plasmid pSCL

C/Species: *Streptomyces clavuligerus*

C/Date: 30-Apr-1998 #sequence_revision 08-May-1998 #text_change 22-Oct-1999

C/Accession: S30401

R:Wu, X.; Roy, K.L.

J. Bacteriol. 175, 37-52, 1993

A:Title: Complete nucleotide sequence of a linear plasmid from *Streptomyces clavuligerus*

A/Reference number: S30400, MUID:99106972, PMID:8416908

A/Accession: S30401

A/Molecule type: DNA

A/Residues: 1-451 <WUX>

A/Cross-references: EMBL:X54107, NID:948758, PIDN:CA38041.1, PID:G581632

C/Genetics:

A/Genome: plasmid pSCL

A/Start codon: GTG

Query Match 37.0%; Score 50; DB 2; Length 451;
Best Local Similarity 52.2%; Pred. No. 19;

Matches 12; Conservative 3; Mismatches 8; Indels 0; Gaps 0;

Qy 2 WLEERDEWVRSIAVDAQHAA 24

Db 220 WRRARIEPRVLLDAENAAQV 242

RESULT 14

C87323

protein-glutamate methyltransferase [imported] - *Caulobacter crescentus*

C/Species: *Caulobacter crescentus*

C/Date: 20-Apr-2001 #sequence_revision 20-Apr-2001 #text_change 15-Sep-2003

C/Accession: C87323

R:Nierman, W.C.; Feldblyum, T.V.; Paulsen, I.T.; Nelson, K.E.; Eisen, J.; Heidelberg, J.I

B.; Laub, M.T.; DeBoy, R.T.; Dodson, R.J.; Durkin, A.S.; Gwinn, M.L.; Haft, D.H.; Kolon

n, J.; Ermolaeva, M.; White, O.; Salzberg, S.L.; Shapiro, L.; Venter, J.C.; Fraser, C.M.

Proc. Natl. Acad. Sci. U.S.A. 98, 4135-4141, 2001

A:Title: Complete Genome Sequence of *Caulobacter crescentus*.

A/Reference number: A87249, MUID:21173698, PMID:11259647

A/Accession: C87323

A/Status: preliminary

A/Molecule type: DNA

A/Residues: 1-347 <STO>

A/Cross-references: GB:AE005673, NID:G13421799, PIDN:AAK22583.1; GSPDB:GN00148

C/Genetics:

A/Gene: CC0597

Query Match 36.3%; Score 49; DB 2; Length 347;
Best Local Similarity 41.7%; Pred. No. 20;

Matches 10; Conservative 5; Mismatches 9; Indels 0; Gaps 0;

Qy 2 WLEERDEWVRSIAVDAQHAA 25

Db 19 WLEAGFIIAIAVDGGQALRKLA 42

RESULT 15

F83397
 Probable peptidase PA1990 [imported] - Pseudomonas aeruginosa (strain PA01)
 C/Species: Pseudomonas aeruginosa
 C/Date: 15-Sep-2000 #sequence_revision 15-Sep-2000 #text_change 31-Dec-2000
 C/Accession: F83397
 R/Stover, C.K.; Pham, X.O.; Erwin, A.L.; Mizoguchi, S.D.; Warren, P.; Hickey, M.J.; B
 adman, S.; Yuan, Y.; Brody, L.L.; Coulter, S.N.; Folger, K.R.; Kas, A.; Lapidis, K.; Lim,
 .; Lory, S.; Olson, M.V.
 Nature 406, 959-964, 2000
 A/Title: Complete genome sequence of Pseudomonas aeruginosa PA01, an opportunistic patho
 A/Reference number: A82950; MUID:20437337; PMID:10984043
 A/Accession: F83397
 A/Status: preliminary
 A/Molecule type: DNA
 A/Residues: 1-608 <STO>
 A/Cross-references: GB:AE004625; GB:AE004091; NID:G947983; PIDN:AAG0378.1; GSPDB:GN001
 A/Experimental source: strain PA01
 C/Genetics:
 A/Gene: PA1990
 C:Superfamily: Synchocystis hypothetical protein slr0825

Query Match 36.3%; Score 49; DB 2; Length 608;
 Best Local Similarity 40.7%; Pred. No. 37;
 Matches 11; Conservative 1; Mismatches 7; Indels 8; Gaps 1;

QY 2 WLEERDEWVRSLA-----VDAQHA 20
 Db 223 WLSPRDGMWRPCAEVALAGRPADNDHA 249

Search completed: September 30, 2004, 11:09:50
 Job time : 18.5905 secs

This Page Blank (uspto)

Fri Oct 1 12:26:17 2004

us-09-988-851a-12.rapb

Page 1

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: September 30, 2004, 11:08:46 ; Search time 65.3714 Seconds
(without alignments)
127.988 Million cell updates/sec

Title: US-09-988-851A-12

Perfect score: 1 TWLEERDEWVRSIAVDAQHARRVAS 26

Sequence: BLOSUM62
Gapop 10.0 , Gapext 0.5

Scoring table: 1351062 segs, 32179191 residues

Searched: Total number of hits satisfying chosen parameters: 1351062

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :

Published Applications AA:*
1: /cgn2_6/ptodata/1/pubppaa/US07_PUBCOMB.pep:*
2: /cgn2_6/ptodata/1/pubppaa/PCR_NEW_PUB.pep:*
3: /cgn2_6/ptodata/1/pubppaa/US06_NEW_PUB.pep:*
4: /cgn2_6/ptodata/1/pubppaa/US06_PUBCOMB.pep:*
5: /cgn2_6/ptodata/1/pubppaa/US07_NEW_PUB.pep:*
6: /cgn2_6/ptodata/1/pubppaa/PCRUS_PUBCOMB.pep:*
7: /cgn2_6/ptodata/1/pubppaa/US08_NEW_PUB.pep:*
8: /cgn2_6/ptodata/1/pubppaa/US08_PUBCOMB.pep:*
9: /cgn2_6/ptodata/1/pubppaa/US08_PUBCOMB.pep:*
10: /cgn2_6/ptodata/1/pubppaa/US09_PUBCOMB.pep:*
11: /cgn2_6/ptodata/1/pubppaa/US09C_PUBCOMB.pep:*
12: /cgn2_6/ptodata/1/pubppaa/US09_NEW_PUB.pep:*
13: /cgn2_6/ptodata/1/pubppaa/US10_PUBCOMB.pep:*
14: /cgn2_6/ptodata/1/pubppaa/US10_PUBCOMB.pep:*
15: /cgn2_6/ptodata/1/pubppaa/US10C_PUBCOMB.pep:*
16: /cgn2_6/ptodata/1/pubppaa/US10_NEW_PUB.pep:*
17: /cgn2_6/ptodata/1/pubppaa/US60_NEW_PUB.pep:*
18: /cgn2_6/ptodata/1/pubppaa/US60_PUBCOMB.pep:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	135	100.0	26	12	US-09-988-851-12
2	97	71.9	30	12	US-09-988-851-11
3	34	40.0	129	12	US-10-282-122A-72804
4	54	40.0	129	12	US-10-282-122A-75294
5	51	37.8	2470	15	US-10-310-154-579
6	51	37.8	2470	15	US-10-369-493-22191
7	51	37.8	2473	15	US-10-369-493-22217
8	50.5	37.4	861	12	US-10-282-122A-43623
9	50	37.0	218	16	US-10-767-701-56064
10	49.5	36.7	151	12	US-10-424-599-181303
11	49.5	36.7	249	12	US-10-424-599-181302
12	49.5	36.7	547	16	US-10-437-963-113173
13	49	36.3	675	14	US-10-156-761-10467
14	49	36.3	871	15	US-10-330-797-3313
15	49	36.3	940	12	US-10-282-122A-63752

16	49	36.3	1503	16	US-10-437-963-164301	Sequence 164301,
17	49	36.3	1524	16	US-10-437-963-164299	Sequence 164299,
18	48	35.6	314	14	US-10-166-225A-86	Sequence 86, Appl
19	48	35.6	385	16	US-10-437-963-164943	Sequence 164943,
20	48	35.6	5317	16	US-10-668-767-59	Sequence 59, Appl
21	47.5	35.2	263	16	US-10-461-194-126	Sequence 126, App
22	47.5	35.2	371	14	US-10-156-761-7595	Sequence 7595, Ap
23	47	34.8	101	14	US-10-190-012-32	Sequence 32, Appl
24	47	34.8	105	10	US-09-922-226-114	Sequence 114, App
25	47	34.8	138	16	US-10-767-701-59340	Sequence 49340, A
26	47	34.8	295	12	US-10-428-114-50264	Sequence 50284, A
27	47	34.8	350	12	US-10-170-385-431	Sequence 431, App
28	47	34.8	350	14	US-10-247-871-143	Sequence 143, App
29	47	34.8	350	16	US-10-755-889-322	Sequence 322, App
30	47	34.8	405	14	US-10-043-487-302	Sequence 302, App
31	47	34.8	465	10	US-09-935-642-3	Sequence 3, Appl1
32	47	34.8	466	9	US-09-923-779-155	Sequence 155, App
33	47	34.8	466	12	US-10-231-956A-497	Sequence 497, App
34	47	34.8	466	12	US-10-231-956A-499	Sequence 499, App
35	47	34.8	466	13	US-10-152-647-3	Sequence 3, Appl1
36	47	34.8	466	13	US-10-152-647-4	Sequence 4, Appl1
37	47	34.8	466	14	US-10-242-943-12	Sequence 12, Appl
38	47	34.8	466	14	US-10-316-253-196	Sequence 196, App
39	47	34.8	466	15	US-10-341-434-107	Sequence 107, App
40	47	34.8	466	16	US-10-408-765A-2034	Sequence 2034, Ap
41	47	34.8	527	12	US-10-425-114-43582	Sequence 43582, A
42	47	34.8	623	12	US-10-425-114-70972	Sequence 70972, A
43	47	34.8	869	16	US-10-437-963-156107	Sequence 156107,
44	46.5	34.4	660	15	US-10-369-493-22892	Sequence 22892, A
45	46.5	34.4	1216	16	US-10-437-963-132249	Sequence 132249,

ALIGNMENTS

RESULT 1
US-09-988-851-12
; Sequence 12, Application US/09998851
; Publication No. US20020034609A1
; GENERAL INFORMATION:
; APPLICANT: Medical Research Council
; APPLICANT: Marsden, Howard
; APPLICANT: Stow, Nigel
; APPLICANT: McLean, Gordon
; TITLE OF INVENTION: ASSAY FOR SCREENING FOR AN ANTI-VIRAL AGENT
; FILE REFERENCE: P29347A
; CURRENT APPLICATION NUMBER: US/09/998,851
; CURRENT FILING DATE: 2002-03-11
; NUMBER OF SEQ ID NOS: 18
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 12
; LENGTH: 26
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Peptide derived from herpes simplex virus.
US-09-988-851-12

Query Match 100.0% Score 135; DB 12; Length 26;
Best Local Similarity 100.0% Pred. No. 1.2e-12;
Matches 26; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 TWLEERDEWVRSIAVDAQHARRVAS 26
Db 1 TWLEERDEWVRSIAVDAQHARRVAS 26
RESULT 2
US-09-988-851-11
; Sequence 11, Application US/09998851
; Publication No. US20020034609A1
; GENERAL INFORMATION:
; APPLICANT: Medical Research Council

```

; APPLICANT: Marsden, Howard
; APPLICANT: Stow, Nigel
; APPLICANT: McLean, Gordon
; TITLE OF INVENTION: ASSAY FOR SCREENING FOR AN ANTI-VIRAL AGENT
; FILE REFERENCE: P293474
; CURRENT APPLICATION NUMBER: US/09/998,851
; CURRENT FILING DATE: 2002-03-11
; NUMBER OF SEQ ID NOS: 18
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 11
; LENGTH: 30
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: peptide derived from herpes simplex virus.
; US-09-998-851-11

Query Match          71.9%; Score 97; DB 12; Length 30;
Best Local Similarity 95.0%; Pred. No. 4.9e-07;
Matches 19; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Cy 7 DEWRSIAVDQAHARRVVS 26
Db 1 DEWRSIAVDQAHARRVVS 20

RESULT 3
US-10-282-122A-72804
; Sequence 72804, Application US/10282122A
; Publication No. US20040029129A1
; GENERAL INFORMATION:
; APPLICANT: Wang, Liangsu
; APPLICANT: Zamudio, Carlos
; APPLICANT: Malone, Cheryl
; APPLICANT: Haselbeck, Robert
; APPLICANT: Ohlsen, Kari
; APPLICANT: Zyskind, Judith
; APPLICANT: Wall, Daniel
; APPLICANT: Trawick, John
; APPLICANT: Carr, Grant
; APPLICANT: Yamamoto, Robert
; APPLICANT: Forsyth, R.
; APPLICANT: Xu, H.
; TITLE OF INVENTION: Identification of Essential Genes in Microorganisms
; FILE REFERENCE: ELITRA.034A
; CURRENT APPLICATION NUMBER: US/10/282,122A
; CURRENT FILING DATE: 2003-02-20
; PRIOR APPLICATION NUMBER: 60/191,078
; PRIOR FILING DATE: 2000-03-21
; PRIOR APPLICATION NUMBER: 60/206,848
; PRIOR FILING DATE: 2000-05-23
; PRIOR APPLICATION NUMBER: 60/207,727
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: 60/230,335
; PRIOR FILING DATE: 2000-09-06
; PRIOR APPLICATION NUMBER: 60/230,347
; PRIOR FILING DATE: 2000-09-09
; PRIOR APPLICATION NUMBER: 60/242,578
; PRIOR FILING DATE: 2000-10-23
; PRIOR APPLICATION NUMBER: 60/253,625
; PRIOR FILING DATE: 2000-11-27
; PRIOR APPLICATION NUMBER: 60/257,931
; PRIOR FILING DATE: 2000-12-22
; PRIOR APPLICATION NUMBER: 60/267,636
; PRIOR FILING DATE: 2001-02-09
; PRIOR APPLICATION NUMBER: 60/269,308
; PRIOR FILING DATE: 2001-02-16
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 78614
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 72804
; LENGTH: 129
; TYPE: PRT
```

```

; ORGANISM: Salmonella paratyphi A
; US-10-282-122A-72804

Query Match          40.0%; Score 54; DB 12; Length 129;
Best Local Similarity 50.0%; Pred. No. 4.5;
Matches 9; Conservative 3; Mismatches 6; Indels 0; Gaps 0;

Cy 2 WLEERDEWRSIAVDQAH 19
Db 37 WLEERDSWKGDIALNGQY 54

RESULT 4
US-10-282-122A-75294
; Sequence 75294, Application US/10282122A
; Publication No. US20040029129A1
; GENERAL INFORMATION:
; APPLICANT: Wang, Liangsu
; APPLICANT: Zamudio, Carlos
; APPLICANT: Malone, Cheryl
; APPLICANT: Haselbeck, Robert
; APPLICANT: Ohlsen, Kari
; APPLICANT: Zyskind, Judith
; APPLICANT: Wall, Daniel
; APPLICANT: Trawick, John
; APPLICANT: Carr, Grant
; APPLICANT: Yamamoto, Robert
; APPLICANT: Forsyth, R.
; APPLICANT: Xu, H.
; TITLE OF INVENTION: Identification of Essential Genes in Microorganisms
; FILE REFERENCE: ELITRA.034A
; CURRENT APPLICATION NUMBER: US/10/282,122A
; CURRENT FILING DATE: 2003-02-20
; PRIOR APPLICATION NUMBER: 60/191,078
; PRIOR FILING DATE: 2000-03-21
; PRIOR APPLICATION NUMBER: 60/206,848
; PRIOR FILING DATE: 2000-05-23
; PRIOR APPLICATION NUMBER: 60/207,727
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: 60/230,335
; PRIOR FILING DATE: 2000-09-06
; PRIOR APPLICATION NUMBER: 60/230,347
; PRIOR FILING DATE: 2000-09-09
; PRIOR APPLICATION NUMBER: 60/242,578
; PRIOR FILING DATE: 2000-10-23
; PRIOR APPLICATION NUMBER: 60/253,625
; PRIOR FILING DATE: 2000-11-27
; PRIOR APPLICATION NUMBER: 60/257,931
; PRIOR FILING DATE: 2000-12-22
; PRIOR APPLICATION NUMBER: 60/267,636
; PRIOR FILING DATE: 2001-02-09
; PRIOR APPLICATION NUMBER: 60/269,308
; PRIOR FILING DATE: 2001-02-16
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 78614
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 75294
; LENGTH: 129
; TYPE: PRT
; ORGANISM: Salmonella typhimurium
; US-10-282-122A-75294

Query Match          40.0%; Score 54; DB 12; Length 129;
Best Local Similarity 50.0%; Pred. No. 4.5;
Matches 9; Conservative 3; Mismatches 6; Indels 0; Gaps 0;

Cy 2 WLEERDEWRSIAVDQAH 19
Db 37 WLEERDSWKGDIALNGQY 54

RESULT 5
US-10-310-154-579
```

```

; Sequence 579, Application US/10310154
; Publication No. US20030233670A1
; GENERAL INFORMATION:
; APPLICANT: Edgerton, Michael D.
; APPLICANT: Chomet, Paul S.
; APPLICANT: Adams, Thomas H.
; APPLICANT: Ruff, Thomas G.
; APPLICANT: Agarwal, Ameeta K.
; APPLICANT: Ahrens, Jeffrey E.
; APPLICANT: Ball, James A.
; APPLICANT: Barni, G.
; APPLICANT: Bell, Erin
; APPLICANT: Boddupalli, Raghava
; APPLICANT: Deikman, Jill
; APPLICANT: Deng, Molian
; APPLICANT: Dong, Jizhuo
; APPLICANT: Duff, Stephen M.
; APPLICANT: Galligan, Meghan M.
; APPLICANT: Hinchey, Brenda S.
; APPLICANT: Huang, Shihshien
; APPLICANT: Johnson, G. Richard
; APPLICANT: Jung, Vincent
; APPLICANT: Kretzmer, Keith A.
; APPLICANT: Laccetti, Lucille B.
; APPLICANT: Lai, Chao-Qiang
; APPLICANT: Lee, Gary
; APPLICANT: Lin, Jie-Yi
; APPLICANT: Liu, Jingdong
; APPLICANT: Lu, Bin
; APPLICANT: Luechy, Michael M.
; APPLICANT: Lund, Adrian
; APPLICANT: Madison, Linda L.
; APPLICANT: Malloy, Kathleen A.
; APPLICANT: McKel, Christine L.
; APPLICANT: Miller, Philip W.
; APPLICANT: Padmavathi, ManchiKant
; APPLICANT: Parnell, Laurence D.
; APPLICANT: Start, William G.
; APPLICANT: Temesen, Dan
; APPLICANT: Vidya, K.R.
; APPLICANT: Wang, Haiyun
; APPLICANT: Xin, Zhanguo
; APPLICANT: Xu, Nanfei
; APPLICANT: Yang, Chunzhi
; APPLICANT: Zeng, Xiaoping
; APPLICANT: Zhang, Qiang
; APPLICANT: Zhao, Yajuan
; APPLICANT: Zhou, Li
; TITLE OF INVENTION: Gene Sequences and Uses Thereof in Plants
; FILE REFERENCE: 38-15(52796)B
; CURRENT APPLICATION NUMBER: US/10/310,154
; CURRENT FILING DATE: 2002-12-04
; PRIOR APPLICATION NUMBER: 60/337,358
; PRIOR FILING DATE: 2001-12-04
; NUMBER OF SEQ ID NOS: 736
; SEQ ID NO 579
; LENGTH: 2470
; TYPE: PRT
; ORGANISM: Saccharomyces cerevisiae
US-10-310-154-579

Query Match          37.8%; Score 51; DB 15; Length 2470;
Best Local Similarity 66.7%; Pred. No. 2.7e+02;
Matches 10; Conservative 1; Mismatches 4; Indels 0; Gaps 0;

```

```

; Publication No. US20030233675A1
; GENERAL INFORMATION:
; APPLICANT: Cao, Yongwei
; APPLICANT: Hinkle, Gregory J.
; APPLICANT: Slater, Steven C.
; APPLICANT: Goldman, Barry S.
; APPLICANT: Chen, Xiandeng
; TITLE OF INVENTION: EXPRESSION OF MICROBIAL PROTEINS IN PLANTS FOR PRODUCTION OF
; FILE REFERENCE: 38-10(52052)B
; CURRENT APPLICATION NUMBER: US/10/369,493
; CURRENT FILING DATE: 2003-02-28
; PRIOR APPLICATION NUMBER: US 60/360,039
; PRIOR FILING DATE: 2002-02-21
; NUMBER OF SEQ ID NOS: 47374
; SEQ ID NO 22191
; LENGTH: 2470
; TYPE: PRT
; ORGANISM: Saccharomyces cerevisiae
US-10-369-493-22191

Query Match          37.8%; Score 51; DB 15; Length 2470;
Best Local Similarity 66.7%; Pred. No. 2.7e+02;
Matches 10; Conservative 1; Mismatches 4; Indels 0; Gaps 0;

; Publication No. US20030233675A1
; GENERAL INFORMATION:
; APPLICANT: Cao, Yongwei
; APPLICANT: Hinkle, Gregory J.
; APPLICANT: Slater, Steven C.
; APPLICANT: Goldman, Barry S.
; APPLICANT: Chen, Xiandeng
; TITLE OF INVENTION: EXPRESSION OF MICROBIAL PROTEINS IN PLANTS FOR PRODUCTION OF
; FILE REFERENCE: 38-10(52052)B
; CURRENT APPLICATION NUMBER: US/10/369,493
; CURRENT FILING DATE: 2003-02-28
; PRIOR APPLICATION NUMBER: US 60/360,039
; PRIOR FILING DATE: 2002-02-21
; NUMBER OF SEQ ID NOS: 47374
; SEQ ID NO 22217
; LENGTH: 2473
; TYPE: PRT
; ORGANISM: Saccharomyces cerevisiae
US-10-369-493-22217

Query Match          37.8%; Score 51; DB 15; Length 2473;
Best Local Similarity 66.7%; Pred. No. 2.7e+02;
Matches 10; Conservative 1; Mismatches 4; Indels 0; Gaps 0;

; Publication No. US200400029129A1
; GENERAL INFORMATION:
; APPLICANT: Wang, Liangsu
; APPLICANT: Zamudio, Carlos
; APPLICANT: Malone, Cheryl
; APPLICANT: Haselbeck, Robert
; APPLICANT: Ohlsen, Karl
US-10-282-122A-43623

Query Match          37.8%; Score 51; DB 15; Length 2470;
Best Local Similarity 66.7%; Pred. No. 2.7e+02;
Matches 10; Conservative 1; Mismatches 4; Indels 0; Gaps 0;

```

```

; Sequence 579, Application US/10310154
; Publication No. US20030233670A1
; GENERAL INFORMATION:
; APPLICANT: Edgerton, Michael D.
; APPLICANT: Chomet, Paul S.
; APPLICANT: Adams, Thomas H.
; APPLICANT: Ruff, Thomas G.
; APPLICANT: Agarwal, Ameeta K.
; APPLICANT: Ahrens, Jeffrey E.
; APPLICANT: Ball, James A.
; APPLICANT: Barni, G.
; APPLICANT: Bell, Erin
; APPLICANT: Boddupalli, Raghava
; APPLICANT: Deikman, Jill
; APPLICANT: Deng, Molian
; APPLICANT: Dong, Jizhuo
; APPLICANT: Duff, Stephen M.
; APPLICANT: Galligan, Meghan M.
; APPLICANT: Hinchey, Brenda S.
; APPLICANT: Huang, Shihshien
; APPLICANT: Johnson, G. Richard
; APPLICANT: Jung, Vincent
; APPLICANT: Kretzmer, Keith A.
; APPLICANT: Laccetti, Lucille B.
; APPLICANT: Lai, Chao-Qiang
; APPLICANT: Lee, Gary
; APPLICANT: Lin, Jie-Yi
; APPLICANT: Liu, Jingdong
; APPLICANT: Lu, Bin
; APPLICANT: Luechy, Michael M.
; APPLICANT: Lund, Adrian
; APPLICANT: Madison, Linda L.
; APPLICANT: Malloy, Kathleen A.
; APPLICANT: McKel, Christine L.
; APPLICANT: Miller, Philip W.
; APPLICANT: Padmavathi, ManchiKant
; APPLICANT: Parnell, Laurence D.
; APPLICANT: Start, William G.
; APPLICANT: Temesen, Dan
; APPLICANT: Vidya, K.R.
; APPLICANT: Wang, Haiyun
; APPLICANT: Xin, Zhanguo
; APPLICANT: Xu, Nanfei
; APPLICANT: Yang, Chunzhi
; APPLICANT: Zeng, Xiaoping
; APPLICANT: Zhang, Qiang
; APPLICANT: Zhao, Yajuan
; APPLICANT: Zhou, Li
; TITLE OF INVENTION: Gene Sequences and Uses Thereof in Plants
; FILE REFERENCE: 38-15(52796)B
; CURRENT APPLICATION NUMBER: US/10/310,154
; CURRENT FILING DATE: 2002-12-04
; PRIOR APPLICATION NUMBER: 60/337,358
; PRIOR FILING DATE: 2001-12-04
; NUMBER OF SEQ ID NOS: 736
; SEQ ID NO 579
; LENGTH: 2470
; TYPE: PRT
; ORGANISM: Saccharomyces cerevisiae
US-10-310-154-579

Query Match          37.8%; Score 51; DB 15; Length 2470;
Best Local Similarity 66.7%; Pred. No. 2.7e+02;
Matches 10; Conservative 1; Mismatches 4; Indels 0; Gaps 0;

```

```

; Publication No. US20030233675A1
; GENERAL INFORMATION:
; APPLICANT: Cao, Yongwei
; APPLICANT: Hinkle, Gregory J.
; APPLICANT: Slater, Steven C.
; APPLICANT: Goldman, Barry S.
; APPLICANT: Chen, Xiandeng
; TITLE OF INVENTION: EXPRESSION OF MICROBIAL PROTEINS IN PLANTS FOR PRODUCTION OF
; FILE REFERENCE: 38-10(52052)B
; CURRENT APPLICATION NUMBER: US/10/369,493
; CURRENT FILING DATE: 2003-02-28
; PRIOR APPLICATION NUMBER: US 60/360,039
; PRIOR FILING DATE: 2002-02-21
; NUMBER OF SEQ ID NOS: 47374
; SEQ ID NO 22191
; LENGTH: 2470
; TYPE: PRT
; ORGANISM: Saccharomyces cerevisiae
US-10-369-493-22191

Query Match          37.8%; Score 51; DB 15; Length 2470;
Best Local Similarity 66.7%; Pred. No. 2.7e+02;
Matches 10; Conservative 1; Mismatches 4; Indels 0; Gaps 0;

```

```

APPLICANT: Zykind, Judith
APPLICANT: Wall, Daniel
APPLICANT: Trawick, John
APPLICANT: Carr, Grant
APPLICANT: Yamamoto, Robert
APPLICANT: Forsyth, R.
APPLICANT: Xu, H
TITLE OF INVENTION: Identification of Essential Genes in Microorganisms
FILE REFERENCE: ELITRA.034A
CURRENT APPLICATION NUMBER: US/10/282,122A
CURRENT FILING DATE: 2003-02-20
PRIOR APPLICATION NUMBER: 60/191,078
PRIOR FILING DATE: 2000-03-21
PRIOR APPLICATION NUMBER: 60/206,848
PRIOR FILING DATE: 2000-05-23
PRIOR APPLICATION NUMBER: 60/207,727
PRIOR FILING DATE: 2000-05-26
PRIOR APPLICATION NUMBER: 60/230,335
PRIOR FILING DATE: 2000-09-06
PRIOR APPLICATION NUMBER: 60/230,347
PRIOR FILING DATE: 2000-09-09
PRIOR APPLICATION NUMBER: 60/242,578
PRIOR FILING DATE: 2000-10-23
PRIOR APPLICATION NUMBER: 60/253,625
PRIOR FILING DATE: 2000-11-27
PRIOR APPLICATION NUMBER: 60/257,931
PRIOR FILING DATE: 2000-12-22
PRIOR APPLICATION NUMBER: 60/267,636
PRIOR FILING DATE: 2001-02-09
PRIOR APPLICATION NUMBER: 60/269,308
PRIOR FILING DATE: 2001-02-16
Remaining Prior Application data removed - See File Wrapper or PALM.
NUMBER OF SEQ ID NOS: 78614
SOFTWARE: PatentIn version 3.1
SEQ ID NO 43623
LENGTH: 861
TYPE: PRT
ORGANISM: Pseudomonas aeruginosa
US-10-282-122A-43623

Query Match      37.4%; Score 50.5; DB 12; Length 861;
Best Local Similarity 43.5%; Pred. No. 1.1e+02;
Matches 10; Conservative 3; Mismatches 7; Indels 3; Gaps 1;

QY      2 WLEERDEWV--RSLAVDAQHAA 21
      |||::|||
Db      819 WEEDRDWLGRRKRAEDDPHAA 841

RESULT 9
US-10-767-701-56064
Sequence 56064, Application US/10767701
Publication No. US20040172684A1
GENERAL INFORMATION:
APPLICANT: Kovalic, David K.
APPLICANT: Zhou, Yihua
APPLICANT: Cao, Yongwei
TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
FILE REFERENCE: 38-21(53535)B
CURRENT APPLICATION NUMBER: US/10/767,701
CURRENT FILING DATE: 2004-01-29
NUMBER OF SEQ ID NOS: 63128
SEQ ID NO 56064
LENGTH: 218
TYPE: PRT
ORGANISM: Sorghum bicolor
FEATURE:
OTHER INFORMATION: Clone ID: 30936847.pap
US-10-767-701-56064

Query Match      37.0%; Score 50; DB 16; Length 218;
Best Local Similarity 38.1%; Pred. No. 30;
Matches 10; Conservative 3; Mismatches 7; Indels 3; Gaps 1;
```

```

Matches 8; Conservative 5; Mismatches 8; Indels 0; Gaps 0;

QY      4 EERDEWVRSIAVDQAARV 24
      |||::|||
Db      17 EDSKXERRAITPQHVARI 37

RESULT 10
US-10-424-599-181303
Sequence 181303, Application US/10424599
Publication No. US20040031072A1
GENERAL INFORMATION:
APPLICANT: La Rosa, Thomas J
APPLICANT: Kovalic, David K
APPLICANT: Zhou, Yihua
APPLICANT: Cao, Yongwei
TITLE OF INVENTION: Soy Nucleic Acid Molecules and Other Molecules Associated With
FILE REFERENCE: 38-21(53223)B
CURRENT APPLICATION NUMBER: US/10/424,599
CURRENT FILING DATE: 2003-04-28
NUMBER OF SEQ ID NOS: 285684
SEQ ID NO 181303
LENGTH: 151
TYPE: PRT
ORGANISM: Glycine max
FEATURE:
OTHER INFORMATION: Clone ID: PAT_MRT3847_134731C.1.pap
US-10-424-599-181303

Query Match      36.7%; Score 49.5; DB 12; Length 151;
Best Local Similarity 52.0%; Pred. No. 24;
Matches 13; Conservative 0; Mismatches 11; Indels 1; Gaps 1;

QY      2 WLEERDEWVRSIAVDQAARV 26
      |||::|||
Db      30 WPERADNW-RDGAHVACRVFARV 53

RESULT 11
US-10-424-599-181302
Sequence 181302, Application US/10424599
Publication No. US20040031072A1
GENERAL INFORMATION:
APPLICANT: La Rosa, Thomas J
APPLICANT: Kovalic, David K
APPLICANT: Zhou, Yihua
APPLICANT: Cao, Yongwei
TITLE OF INVENTION: Soy Nucleic Acid Molecules and Other Molecules Associated With
FILE REFERENCE: 38-21(53223)B
CURRENT APPLICATION NUMBER: US/10/424,599
CURRENT FILING DATE: 2003-04-28
NUMBER OF SEQ ID NOS: 285684
SEQ ID NO 181302
LENGTH: 249
TYPE: PRT
ORGANISM: Glycine max
FEATURE:
OTHER INFORMATION: Clone ID: PAT_MRT3847_134730C.1.pap
US-10-424-599-181302

Query Match      36.7%; Score 49.5; DB 12; Length 249;
Best Local Similarity 52.0%; Pred. No. 40;
Matches 13; Conservative 0; Mismatches 11; Indels 1; Gaps 1;

QY      2 WLEERDEWVRSIAVDQAARV 26
      |||::|||
Db      30 WPERADNW-RDGAHVACRVFARV 53

RESULT 12
US-10-437-963-113173
```

```
; Sequence 113173, Application US/10437963
; Publication No. US20040123343A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa, Thomas J.
; APPLICANT: Kovalic, David K.
; APPLICANT: Zhou, Yihua
; APPLICANT: Cao, Yongwei
; APPLICANT: Wu, Wei
; APPLICANT: Boukharov, Andrey A.
; APPLICANT: Barbazuk, Brad
; APPLICANT: Li, Ping
; TITLE OF INVENTION: Rice Nucleic Acid Molecules and Other Molecules Associated With
; FILE REFERENCE: 38-21(53221)B
; CURRENT APPLICATION NUMBER: US/10/437,963
; CURRENT FILING DATE: 2003-05-14
; NUMBER OF SEQ ID NOS: 204966
; SEQ ID NO 113173
; LENGTH: 547
; TYPE: PRT
; ORGANISM: Oryza sativa
; FEATURE:
; OTHER INFORMATION: Clone ID: PAT_MRT4530_16987C.1 pep
; US-10-437-963-113173
```

```
Query Match 36.7%; Score 49.5; DB 16; Length 547;
Best Local Similarity 51.9%; Pred. No. 92;
Matches 14; Conservative 2; Mismatches 4; Indels 7; Gaps 2;
```

```
Oy 2 TWLEPDEWVRSIAVDAQH---AARV 24
Db 269 WLNED--RSKALDEKISRGDAAARV 292
```

```
RESULT 13
; US-10-156-761-10467
; Sequence 10467, Application US/10156761
; Publication No. US20030119018A1
; GENERAL INFORMATION:
; APPLICANT: OMURA, SATOSHI
; APPLICANT: IKEDA, HARUO
; APPLICANT: ISHIKAWA, JUN
; APPLICANT: HORIKAWA, HIROSHI
; APPLICANT: SHIBA, TADAYOSHI
; APPLICANT: SAKAKI, YOSHIYUKI
; APPLICANT: HATTORI, MASAHIRA
; TITLE OF INVENTION: NOVEL POLYNUCLEOTIDES
; FILE REFERENCE: 249-262
; CURRENT APPLICATION NUMBER: US/10/156,761
; CURRENT FILING DATE: 2002-05-29
; PRIOR APPLICATION NUMBER: JP 2001-204089
; PRIOR FILING DATE: 2001-05-30
; PRIOR APPLICATION NUMBER: JP 2001-272697
; PRIOR FILING DATE: 2001-08-02
; NUMBER OF SEQ ID NOS: 15109
; SEQ ID NO 10467
; LENGTH: 675
; TYPE: PRT
; ORGANISM: Streptomyces avermitilis
; US-10-156-761-10467
```

```
Query Match 36.3%; Score 49; DB 14; Length 675;
Best Local Similarity 55.6%; Pred. No. 1.4e+02;
Matches 10; Conservative 1; Mismatches 7; Indels 0; Gaps 0;
```

```
Oy 3 LEERDEWVRSIAVDAQH 20
Db 341 LEERDRMLELMPDTRA 358
```

```
RESULT 14
; US-10-320-797-3313
; Sequence 3313, Application US/10320797
```

```
; Publication No. US20040014955A1
; GENERAL INFORMATION:
; APPLICANT: Eroshtkin, Alexey M.
; APPLICANT: Zamudio, Carlos
; TITLE OF INVENTION: IDENTIFICATION OF ESSENTIAL GENES OF CRYPTOCOCCUS NEOFORMANS AND
; FILE REFERENCE: 10182-021-999
; CURRENT APPLICATION NUMBER: US/10/320,797
; CURRENT FILING DATE: 2002-12-16
; PRIOR APPLICATION NUMBER: 60/341,261
; PRIOR FILING DATE: 2001-12-17
; NUMBER OF SEQ ID NOS: 3361
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 3313
; LENGTH: 871
; TYPE: PRT
; ORGANISM: Cryptococcus neoformans
; US-10-320-797-3313
```

```
Query Match 36.3%; Score 49; DB 15; Length 871;
Best Local Similarity 38.5%; Pred. No. 1.8e+02;
Matches 10; Conservative 4; Mismatches 12; Indels 0; Gaps 0;
```

```
Oy 1 TWLEPDEWVRSIAVDAQHARRVAS 26
Db 828 TWRAELGEWTVRGVDAQNIGSEKAT 853
```

```
RESULT 15
; US-10-282-122A-63752
; Sequence 63752, Application US/10282122A
; Publication No. US20040029129A1
; GENERAL INFORMATION:
; APPLICANT: Wang, Liangsu
; APPLICANT: Zamudio, Carlos
; APPLICANT: Malone, Cheryl
; APPLICANT: Haselbeck, Robert
; APPLICANT: Ohlsen, Karl
; APPLICANT: Zyskind, Judith
; APPLICANT: Wall, Daniel
; APPLICANT: Trawick, John
; APPLICANT: Carr, Grant
; APPLICANT: Yamamoto, Robert
; APPLICANT: Forsyth, R.
; TITLE OF INVENTION: Identification of Essential Genes in Microorganisms
; FILE REFERENCE: ELITRA.034A
; CURRENT APPLICATION NUMBER: US/10/282,122A
; CURRENT FILING DATE: 2003-02-20
; PRIOR APPLICATION NUMBER: 60/191,078
; PRIOR FILING DATE: 2000-03-21
; PRIOR APPLICATION NUMBER: 60/206,848
; PRIOR FILING DATE: 2000-05-23
; PRIOR APPLICATION NUMBER: 60/207,727
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: 60/230,335
; PRIOR FILING DATE: 2000-09-06
; PRIOR APPLICATION NUMBER: 60/230,347
; PRIOR FILING DATE: 2000-09-09
; PRIOR APPLICATION NUMBER: 60/242,578
; PRIOR FILING DATE: 2000-10-23
; PRIOR APPLICATION NUMBER: 60/253,625
; PRIOR FILING DATE: 2000-11-27
; PRIOR APPLICATION NUMBER: 60/257,931
; PRIOR FILING DATE: 2000-12-22
; PRIOR APPLICATION NUMBER: 60/267,636
; PRIOR FILING DATE: 2001-02-09
; PRIOR APPLICATION NUMBER: 60/269,308
; PRIOR FILING DATE: 2001-02-16
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 78614
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 63752
```

; LENGTH: 940
; TYPE: PRT
; ORGANISM: Mycobacterium leprae
US-10-282-122A-63752

Query Match 36.3%; Score 49; DB 12; Length 940;
Best Local Similarity 60.0%; Pred. No. 1.9e+02;
Matches 9; Conservative 3; Mismatches 3; Indels 0; Gaps 0;

QY 6 RDEWVRSLAVDAQHA 20
|||:|:|||||
Db 732 RDELEALIDDAQHA 746

Search completed: September 30, 2004, 11:33:36
Job time : 66.3714 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Comogen Ltd.

OM protein - protein search, using sw model
Run on: September 30, 2004, 10:58:43 ; Search time 18.0762 Seconds
(Without alignments)
74.257 Million cell updates/sec

Title: US-09-988-851A-12
Perfect score: 135
Sequence: 1 TWLEERDEWVRSIAYDAQHARRVAS 26
Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 389414 seqs, 51625971 residues
Total number of hits satisfying chosen parameters: 389414

Minimum DB seq length: 0
Maximum DB seq length: 200000000
Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Issued Patents AA:*
1: /cgn2_6/ptodata/2/1aa/5A.COMB.pep:*
2: /cgn2_6/ptodata/2/1aa/5B.COMB.pep:*
3: /cgn2_6/ptodata/2/1aa/5A.COMB.pep:*
4: /cgn2_6/ptodata/2/1aa/5B.COMB.pep:*
5: /cgn2_6/ptodata/2/1aa/PTCUS.COMB.pep:*
6: /cgn2_6/ptodata/2/1aa/backfile1.pep:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	135	100.0	26	4	US-09-230-405-12
2	97	71.9	30	4	US-09-230-405-11
3	51	37.8	2470	4	US-08-265-967C-2
4	51	37.8	2470	4	US-08-305-7908-3
5	51	37.8	2474	4	US-08-265-967C-3
6	51	37.8	2474	4	US-08-305-7908-4
7	50.5	37.4	555	4	US-09-252-991A-25453
8	50.5	37.4	871	4	US-09-252-991A-33057
9	50	37.0	335	4	US-09-252-991A-25335
10	50	37.0	420	4	US-09-489-039A-7218
11	49	36.3	614	4	US-09-252-991A-20060
12	47	34.8	96	3	US-08-407-165-4
13	47	34.8	103	4	US-09-511-861A-17
14	47	34.8	466	4	US-09-610-401-3
15	47	34.8	466	4	US-09-610-401-4
16	47	34.8	466	4	US-09-167-206-12
17	46	34.1	298	4	US-09-328-352-6225
18	46	34.1	424	4	US-09-252-991A-26486
19	45.5	33.7	263	4	US-08-708-573F-10
20	45.5	33.7	268	4	US-08-708-573F-14
21	45.5	33.7	399	2	US-08-833-581A-2
22	45.5	33.7	399	2	US-09-023-591A-2
23	45	33.3	31	4	US-09-228-990-68
24	45	33.3	124	4	US-09-511-861A-16
25	45	33.3	561	4	US-09-489-039A-8807
26	45	33.3	742	4	US-09-252-991A-32049
27	45	33.3	977	4	US-09-302-812-2

28	45	33.3	977	4	US-09-511-477-2	Sequence 2, Appli
29	45	33.3	977	4	US-09-511-507-2	Sequence 2, Appli
30	44	32.6	257	4	US-09-252-991A-21740	Sequence 21740, A
31	44	32.6	383	4	US-09-491-577-18	Sequence 18, Appl
32	44	32.6	394	2	US-08-833-581A-4	Sequence 4, Appli
33	44	32.6	394	3	US-09-023-591A-2	Sequence 4, Appli
34	44	32.6	843	4	US-09-252-991A-32609	Sequence 32609, A
35	43	31.9	93	4	US-09-252-991A-31966	Sequence 31966, A
36	43	31.9	332	4	US-09-543-681A-6870	Sequence 6870, Ap
37	43	31.9	552	4	US-09-907-794A-170	Sequence 170, App
38	43	31.9	552	4	US-09-905-125A-170	Sequence 170, App
39	43	31.9	552	4	US-09-902-775A-170	Sequence 170, App
40	43	31.9	561	4	US-09-252-991A-26778	Sequence 26778, A
41	43	31.9	562	3	US-09-012-515A-14	Sequence 14, Appl
42	43	31.9	562	3	US-08-360-144A-14	Sequence 14, Appl
43	43	31.9	562	4	US-09-012-504A-14	Sequence 14, Appl
44	43	31.9	562	4	US-09-012-399A-14	Sequence 14, Appl
45	43	31.9	562	5	PCR-US95-06722-14	Sequence 14, Appl

ALIGNMENTS

RESULT 1
US-09-230-405-12
Sequence 12, Application US/09230405
Patent No. 6337074
GENERAL INFORMATION:
APPLICANT: Medical Research Council
TITLE OF INVENTION: ANTI-VIRAL AGENT
FILE REFERENCE: P17970C
CURRENT APPLICATION NUMBER: US/09/230,405
CURRENT FILING DATE: 1999-01-25
NUMBER OF SEQ ID NOS: 18
SOFTWARE: Patent In Ver. 2.1
SEQ ID NO 12
LENGTH: 26
TYPE: PRT
ORGANISM: Artificial Sequence
FEATURE:
OTHER INFORMATION: Description of Artificial Sequence:peptide derived
US-09-230-405-12
Query Match 100.0%; Score 135; DB 4; Length 26;
Best Local Similarity 100.0%; Pred. No. 4.7e-14;
Matches 26; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 1 TWLEERDEWVRSIAYDAQHARRVAS 26
Db 1 TWLEERDEWVRSIAYDAQHARRVAS 26
RESULT 2
US-09-230-405-11
Sequence 11, Application US/09230405
Patent No. 6337074
GENERAL INFORMATION:
APPLICANT: Medical Research Council
TITLE OF INVENTION: ANTI-VIRAL AGENT
FILE REFERENCE: P17970C
CURRENT APPLICATION NUMBER: US/09/230,405
CURRENT FILING DATE: 1999-01-25
NUMBER OF SEQ ID NOS: 18
SOFTWARE: Patent In Ver. 2.1
SEQ ID NO 11
LENGTH: 30
TYPE: PRT
ORGANISM: Artificial Sequence
FEATURE:
OTHER INFORMATION: Description of Artificial Sequence:peptide derived
US-09-230-405-11

Query Match 71.9%; Score 97; DB 4; Length 30;
 Best Local Similarity 95.0%; Pred. No. 3,36-08;
 Matches 19; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy 7 DEWYS:AVDAQAHARRVAS 26
 |||||
 Db 1 DEWYSLAVDAQAHARRVAS 20

RESULT 3

US-08-265-967C-2
 ; Sequence 2, Application US/08265967C
 ; Patent No. 6476200

; GENERAL INFORMATION:

; APPLICANT: SABATINI, DAVID M.
 ; APPLICANT: ERDUMERT-BROWAGE, HEDIYE

; APPLICANT: LUI, MARY

; APPLICANT: TEMPEST, PAUL

; TITLE OF INVENTION: MAMMALIAN PROTEINS THAT BIND TO FKBP12

; TITLE OF INVENTION: IN A RAPAMYCIN-DEPENDENT FASHION

; NUMBER OF SEQUENCES: 14

; CORRESPONDENCE ADDRESS:

; ADDRESSEE: BANNER & ALLEGRETTI, LTD

; STREET: 1001 G STREET, N.W., 11TH FLOOR

; CITY: WASHINGTON

; STATE: D.C.

; COUNTRY: U.S.A.

; ZIP: 20001-4597

; COMPUTER READABLE FORM:

; MEDIUM TYPE: Floppy disk

; OPERATING SYSTEM: IBM PC compatible

; SOFTWARE: Patent Release #1.0, Version #1.25

; CURRENT APPLICATION DATA:

; APPLICATION NUMBER: US/08/265,967C

; FILING DATE: 27-JUN-1994

; CLASSIFICATION: 435

; ATTORNEY/AGENT INFORMATION:

; NAME: KAGAN, SARAH A.

; REGISTRATION NUMBER: 32,141

; REFERENCE/DOCKET NUMBER: 01107.46363

; TELECOMMUNICATION INFORMATION:

; TELEPHONE: 202-508-9100

; TELEFAX: 202-508-9299

; TELEX: 197430 BMB UT

; INFORMATION FOR SEQ ID NO: 2:

; SEQUENCE CHARACTERISTICS:

; LENGTH: 2470 amino acids

; TYPE: amino acid

; TOPOLOGY: linear

; MOLECULE TYPE: protein

; ORIGINAL SOURCE:

; ORGANISM: Saccharomyces cerevisiae

US-08-265-967C-2

Query Match 37.8%; Score 51; DB 4; Length 2470;
 Best Local Similarity 66.7%; Pred. No. 53;
 Matches 10; Conservative 1; Mismatches 4; Indels 0; Gaps 0;

Qy 1 TWLEERDEWYSLAV 15
 |||||
 Db 2249 TWLEERTTYRSLAV 2263

RESULT 4

US-08-305-790B-3

; Sequence 3, Application US/08305790B

; Patent No. 6492106

; GENERAL INFORMATION:

; APPLICANT: SABATINI, DAVID M.

; APPLICANT: ERDUMERT-BROWAGE, HEDIYE

; APPLICANT: LUI, MARY

; APPLICANT: TEMPEST, PAUL

; TITLE OF INVENTION: MAMMALIAN PROTEINS THAT BIND TO FKBP12

; TITLE OF INVENTION: IN A RAPAMYCIN-DEPENDENT FASHION

; NUMBER OF SEQUENCES: 14

; CORRESPONDENCE ADDRESS:

; ADDRESSEE: BANNER & ALLEGRETTI, LTD

; STREET: 1001 G STREET, N.W., 11TH FLOOR

; CITY: WASHINGTON

; STATE: D.C.

; COUNTRY: U.S.A.

; ZIP: 20001-4597

; COMPUTER READABLE FORM:

APPLICANT: LUI, MARY
 APPLICANT: TEMPEST, PAUL
 APPLICANT: SNYDER, SOLOMON H.
 TITLE OF INVENTION: MAMMALIAN PROTEINS THAT BIND TO FKBP12
 TITLE OF INVENTION: IN A RAPAMYCIN-DEPENDENT FASHION
 NUMBER OF SEQUENCES: 15
 CORRESPONDENCE ADDRESS:
 ADDRESSEE: BANNER & ALLEGRETTI, LTD
 STREET: 1001 G STREET, N.W., 11TH FLOOR
 CITY: WASHINGTON
 STATE: D.C.
 COUNTRY: U.S.A.
 ZIP: 20001-4597

COMPUTER READABLE FORM:
 MEDIUM TYPE: Floppy disk
 OPERATING SYSTEM: IBM PC compatible
 SOFTWARE: Patent Release #1.0, Version #1.25
 CURRENT APPLICATION DATA:
 APPLICATION NUMBER: US/08/305,790B
 FILING DATE:
 CLASSIFICATION: 536
 PRIOR APPLICATION DATA:
 APPLICATION NUMBER: US 08/265,967
 FILING DATE: 27-JUN-1994
 ATTORNEY/AGENT INFORMATION:
 NAME: KAGAN, SARAH A.
 REGISTRATION NUMBER: 32,141
 REFERENCE/DOCKET NUMBER: 01107.47225
 TELECOMMUNICATION INFORMATION:
 TELEPHONE: 202-508-9100
 TELEFAX: 202-508-9299
 TELEX: 197430 BMB UT
 INFORMATION FOR SEQ ID NO: 3:
 SEQUENCE CHARACTERISTICS:
 LENGTH: 2470 amino acids
 TYPE: amino acid
 TOPOLOGY: linear
 MOLECULE TYPE: protein
 ORIGINAL SOURCE:
 ORGANISM: Saccharomyces cerevisiae

US-08-305-790B-3

Query Match 37.8%; Score 51; DB 4; Length 2470;
 Best Local Similarity 66.7%; Pred. No. 53;
 Matches 10; Conservative 1; Mismatches 4; Indels 0; Gaps 0;

Qy 1 TWLEERDEWYSLAV 15
 |||||
 Db 2249 TWLEERTTYRSLAV 2263

RESULT 5

US-08-265-967C-3

; Sequence 3, Application US/08265967C

; Patent No. 6476200

; GENERAL INFORMATION:

; APPLICANT: SABATINI, DAVID M.

; APPLICANT: ERDUMERT-BROWAGE, HEDIYE

; APPLICANT: LUI, MARY

; APPLICANT: TEMPEST, PAUL

; TITLE OF INVENTION: MAMMALIAN PROTEINS THAT BIND TO FKBP12

; TITLE OF INVENTION: IN A RAPAMYCIN-DEPENDENT FASHION

; NUMBER OF SEQUENCES: 14

; CORRESPONDENCE ADDRESS:

; ADDRESSEE: BANNER & ALLEGRETTI, LTD

; STREET: 1001 G STREET, N.W., 11TH FLOOR

; CITY: WASHINGTON

; STATE: D.C.

; COUNTRY: U.S.A.

; ZIP: 20001-4597

; COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk
 COMPUTER: IBM PC compatible
 OPERATING SYSTEM: PC-DOS/MS-DOS
 SOFTWARE: Patentin Release #1.0, Version #1.25
 CURRENT APPLICATION DATA:
 APPLICATION NUMBER: US/08/265,967C
 FILING DATE: 27-JUN-1994
 CLASSIFICATION: 435
 ATTORNEY/AGENT INFORMATION:
 NAME: KAGAN, SARAH A.
 REGISTRATION NUMBER: 32,141
 REFERENCE/DOCKET NUMBER: 01107,46363
 TELECOMMUNICATION INFORMATION:
 TELEPHONE: 202-508-9100
 TELEFAX: 202-508-9299
 TELEX: 197430 BMBB UT
 INFORMATION FOR SEQ ID NO: 3:
 SEQUENCE CHARACTERISTICS:
 LENGTH: 2474 amino acids
 TYPE: amino acid
 TOPOLOGY: linear
 MOLECULE TYPE: protein
 ORIGINAL SOURCE:
 ORGANISM: *Saccharomyces cerevisiae*
 US-08-265-967C-3

Query Match 37.8%; Score 51; DB 4; Length 2474;
 Best Local Similarity 66.7%; Pred. No. 53;
 Matches 10; Conservative 1; Mismatches 4; Indels 0; Gaps 0;

QY 1 TWLEERDEWVSLAV 15
 DB 2253 TWLEERTTYRSLAV 2267

RESULT 6
 US-08-305-790B-4
 Sequence 4, Application US/08305790B
 Patent No. 6492106
 GENERAL INFORMATION:
 APPLICANT: SABATINI, DAVID M.
 APPLICANT: ERDJUMENT-BRONAGE, HEDIYE
 APPLICANT: LUI, MARY
 APPLICANT: TEMPEST, PAUL
 APPLICANT: SNYDER, SOLOMON H.
 TITLE OF INVENTION: MAMMALIAN PROTEINS THAT BIND TO FKBP12
 TITLE OF INVENTION: IN A RAPAMYCIN-DEPENDENT FASHION
 NUMBER OF SEQUENCES: 15
 CORRESPONDENCE ADDRESS:
 ADDRESSEE: BANNER & ALLEGRETTI, LTD
 STREET: 1001 G STREET, N.W., 11TH FLOOR
 CITY: WASHINGTON
 STATE: D.C.
 COUNTRY: U.S.A.
 ZIP: 20001-4597
 COMPUTER READABLE FORM:
 MEDIUM TYPE: Floppy disk
 COMPUTER: IBM PC compatible
 OPERATING SYSTEM: PC-DOS/MS-DOS
 SOFTWARE: Patentin Release #1.0, Version #1.25
 CURRENT APPLICATION DATA:
 APPLICATION NUMBER: US/08/305,790B
 FILING DATE:
 CLASSIFICATION: 536
 PRIOR APPLICATION DATA:
 APPLICATION NUMBER: US 08/265,967
 FILING DATE: 27-JUN-1994
 ATTORNEY/AGENT INFORMATION:
 NAME: KAGAN, SARAH A.
 REGISTRATION NUMBER: 32,141
 REFERENCE/DOCKET NUMBER: 01107,47225
 TELECOMMUNICATION INFORMATION:
 TELEPHONE: 202-508-9100

TELEFAX: 202-508-9299
 TELEX: 197430 BMBB UT
 INFORMATION FOR SEQ ID NO: 4:
 SEQUENCE CHARACTERISTICS:
 LENGTH: 2474 amino acids
 TYPE: amino acid
 TOPOLOGY: linear
 MOLECULE TYPE: protein
 ORIGINAL SOURCE:
 ORGANISM: *Saccharomyces cerevisiae*
 US-08-305-790B-4
 Query Match 37.8%; Score 51; DB 4; Length 2474;
 Best Local Similarity 66.7%; Pred. No. 53;
 Matches 10; Conservative 1; Mismatches 4; Indels 0; Gaps 0;

QY 1 TWLEERDEWVSLAV 15
 DB 2253 TWLEERTTYRSLAV 2267

RESULT 7
 US-09-252-991A-25453
 Sequence 25453, Application US/09252991A
 Patent No. 6551795
 GENERAL INFORMATION:
 APPLICANT: Marc J. Rubenfield et al.
 TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
 TITLE OF INVENTION: AERUGINOSA FOR DIAGNOSTICS AND THERAPEUTICS
 FILE REFERENCE: 107196,136
 CURRENT APPLICATION NUMBER: US/09/252,991A
 CURRENT FILING DATE: 1999-02-18
 PRIOR FILING DATE: 1998-02-18
 PRIOR APPLICATION NUMBER: US 60/094,190
 PRIOR FILING DATE: 1998-07-27
 NUMBER OF SEQ ID NOS: 33142
 SEQ ID NO 25453
 LENGTH: 555
 TYPE: PRT
 ORGANISM: *Pseudomonas aeruginosa*
 US-09-252-991A-25453

Query Match 37.4%; Score 50.5; DB 4; Length 555;
 Best Local Similarity 50.0%; Pred. No. 11;
 Matches 14; Conservative 0; Mismatches 5; Indels 9; Gaps 2;

QY 2 WLEER-----DEWVSLAVDAQHAAARV 24
 DB 485 WLEERPVLVDEW-----AADDPAFRRV 508

RESULT 8
 US-09-252-991A-33057
 Sequence 33057, Application US/09252991A
 Patent No. 6551795
 GENERAL INFORMATION:
 APPLICANT: Marc J. Rubenfield et al.
 TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
 TITLE OF INVENTION: AERUGINOSA FOR DIAGNOSTICS AND THERAPEUTICS
 FILE REFERENCE: 107196,136
 CURRENT APPLICATION NUMBER: US/09/252,991A
 CURRENT FILING DATE: 1999-02-18
 PRIOR FILING DATE: 1998-02-18
 PRIOR APPLICATION NUMBER: US 60/074,788
 PRIOR FILING DATE: 1998-07-27
 PRIOR APPLICATION NUMBER: US 60/094,190
 PRIOR FILING DATE: 1998-07-27
 NUMBER OF SEQ ID NOS: 33142
 SEQ ID NO 33057
 LENGTH: 871
 TYPE: PRT
 ORGANISM: *Pseudomonas aeruginosa*
 US-09-252-991A-33057

Query Match 37.4%; Score 50.5; DB 4; Length 871;
Best Local Similarity 43.5%; Pred. No. 19;
Matches 10; Conservative 3; Mismatches 7; Indels 3; Gaps 1;
QY 2 WLEERDEWV---RSLAVDAQHAA 21
DB 829 WEEDRDMDLGMWRKAEADDPHAA 851

RESULT 9
US-09-252-991A-25335
; Sequence 25335, Application US/09252991A
; Patent No. 6551795
; GENERAL INFORMATION:
; APPLICANT: Marc J. Rubenfield et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
; FILE REFERENCE: 107196.136
; CURRENT APPLICATION NUMBER: US/09/252,991A
; PRIOR FILING DATE: 1999-02-18
; PRIOR APPLICATION NUMBER: US 60/074,788
; PRIOR FILING DATE: 1998-02-18
; PRIOR APPLICATION NUMBER: US 60/094,190
; PRIOR FILING DATE: 1998-07-27
; NUMBER OF SEQ ID NOS: 33142
; SEQ ID NO 25335
; LENGTH: 335
; TYPE: PRT
; ORGANISM: Pseudomonas aeruginosa
US-09-252-991A-25335

Query Match 37.0%; Score 50; DB 4; Length 335;
Best Local Similarity 61.1%; Pred. No. 7.5;
Matches 11; Conservative 2; Mismatches 5; Indels 0; Gaps 0;
QY 4 EERDEWVSLAVDAQHAA 21
DB 35 EERDEFORMLADPRHAA 52

RESULT 10
US-09-489-039A-7218
; Sequence 7218, Application US/09489039A
; Patent No. 6610836
; GENERAL INFORMATION:
; APPLICANT: Gary Breton et. al
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO KLEBSIELLA
; FILE REFERENCE: 2709.2004001
; CURRENT APPLICATION NUMBER: US/09/489,039A
; PRIOR FILING DATE: 2000-01-27
; PRIOR APPLICATION NUMBER: US 60/117,747
; PRIOR FILING DATE: 1999-01-29
; NUMBER OF SEQ ID NOS: 14342
; SEQ ID NO 7218
; LENGTH: 420
; TYPE: PRT
; ORGANISM: Klebsiella pneumoniae
US-09-489-039A-7218

Query Match 37.0%; Score 50; DB 4; Length 420;
Best Local Similarity 45.5%; Pred. No. 9.8;
Matches 10; Conservative 2; Mismatches 10; Indels 0; Gaps 0;
QY 5 EERDEWVSLAVDAQHAA 26
DB 42 ERDDMARQLTALRQMAEQAA 63

RESULT 11
US-09-252-991A-20060
; Sequence 20060, Application US/09252991A

; Patent No. 6551795
; GENERAL INFORMATION:
; APPLICANT: Marc J. Rubenfield et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
; FILE REFERENCE: 107196.136
; CURRENT APPLICATION NUMBER: US/09/252,991A
; PRIOR FILING DATE: 1999-02-18
; PRIOR APPLICATION NUMBER: US 60/074,788
; PRIOR FILING DATE: 1998-02-18
; PRIOR APPLICATION NUMBER: US 60/094,190
; PRIOR FILING DATE: 1998-07-27
; NUMBER OF SEQ ID NOS: 33142
; SEQ ID NO 20060
; LENGTH: 614
; TYPE: PRT
; ORGANISM: Pseudomonas aeruginosa
US-09-252-991A-20060

Query Match 36.3%; Score 49; DB 4; Length 614;
Best Local Similarity 40.7%; Pred. No. 21;
Matches 11; Conservative 1; Mismatches 7; Indels 8; Gaps 1;
QY 2 WLEERDEWVSLA-----VDAQHAA 20
DB 229 WLSDRDGMWRPCAEVALAGRPADHAA 255

RESULT 12
US-08-407-165-4
; Sequence 4, Application US/08407165
; Patent No. 6054280
; GENERAL INFORMATION:
; APPLICANT: LEMMON, MARK A.
; APPLICANT: FERGUSON, KATHRYN M.
; APPLICANT: SIGLER, PAUL B.
; APPLICANT: SCHLESSINGER, JOSEPH
; TITLE OF INVENTION: METHODS FOR DIAGNOSIS AND TREATMENT OF
; FILE REFERENCE: 211/156
; CURRENT APPLICATION NUMBER: US/08/407,165
; PRIOR FILING DATE: 1995-03-20
; NUMBER OF SEQ ID NOS: 4
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 4
; LENGTH: 96
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; OTHER INFORMATION: plectstrin N
US-08-407-165-4

Query Match 34.8%; Score 47; DB 3; Length 96;
Best Local Similarity 66.7%; Pred. No. 5.1;
Matches 8; Conservative 2; Mismatches 2; Indels 0; Gaps 0;
QY 2 WLEERDEWVSL 13
DB 78 FLEERDAWVRI 89

RESULT 13
US-09-511-881A-17
; Sequence 17, Application US/09511881A
; Patent No. 6495674
; GENERAL INFORMATION:
; APPLICANT: SALK INSTITUTE
; APPLICANT: LEMKE, GREGORY
; APPLICANT: NGUYEN, ANDREW
; APPLICANT: KRAPP, RALF
; TITLE OF INVENTION: EYECTING AND THEIR USE
; FILE REFERENCE: SALKINS.001A
; CURRENT APPLICATION NUMBER: US/09/511,881A

```

? CURRENT FILING DATE: 2000-02-25
? NUMBER OF SEQ ID NOS: 17
? SOFTWARE: FastSeq for Windows Version 4.0
? SEQ ID NO: 17
? LENGTH: 103
? TYPE: PRT
? ORGANISM: Homo sapiens
? FEATURE:
? NAME/KEY: DOMAIN
? LOCATION: (1)...(103)
? OTHER INFORMATION: Pleckstrin homology domain
? US-09-511-861A-17

```

Query Match	34.8%	Score 47	DB 4	Length 103
Best Local Similarity	66.7%	Pred. No. 5.5		
Matches	8	Conservative	2	Mismatches 2; Indels 0; Gaps 0;

```
QY      2 WLEERDEWVRL 13
          :||||| :
Db      85 FLEERDAWVRDI 96
```

RESULT 14
 US-09-610-401-3
 ; Sequence 3, Application US/09610401
 ; Patent No. 6417316
 ; GENERAL INFORMATION:
 ; APPLICANT: MORISHIMA, No. 6417336uniro,
 ; APPLICANT: NAKANISHI, Keiko,
 ; APPLICANT: SHIBATA, Takehiko
 ; TITLE OF INVENTION: Antibody against cleavage product of vimentin
 ; FILE REFERENCE: 522-1104
 ; CURRENT APPLICATION NUMBER: US/09/610,401
 ; CURRENT FILING DATE: 2000-07-05
 ; PRIOR APPLICATION NUMBER: JP 11-193235
 ; PRIOR FILING DATE: 1999-07-07
 ; NUMBER OF SEQ ID NOS: 4
 ; SOFTWARE: Patentin Ver. 2.0
 ; SEQ ID NO: 3
 ; LENGTH: 466
 ; TYPE: PR1
 ; ORGANISM: Homo sapiens
 ; US-09-610-401-3

Query Match	34.8%	Score 47;	DB 4;	Length 466;
Best Local Similarity	42.9%;	Pred. No. 31;		
Matches	9;	Conservative	3;	Mismatches 9;
				Indels 0;
				Gaps 0;

```

QY      3 LEERDEWVRS LAVD AQAARR 23
      284 LQEA EFWYKSKFADLSEANR 304
Db

```

RESULT 15
 US-09-610-401-4
 ; Sequence 4, Application US/09610401
 ; Patent No. 6417336
 ; GENERAL INFORMATION:
 ; APPLICANT: MORISHIMA, No. 6417336unhiro,
 ; APPLICANT: NAKANISHI, Keiko,
 ; TITLE OF INVENTION: Antibody against cleavage product of vimentin
 ; FILE REFERENCE: 522.1004
 ; CURRENT APPLICATION NUMBER: US/09/610,401
 ; CURRENT FILING DATE: 2000-07-05
 ; PRIOR APPLICATION NUMBER: JP 11-193235
 ; PRIOR FILING DATE: 1999-07-07
 ; NUMBER OF SEQ ID NOS: 4
 ; SOFTWARE: PatentIn Ver. 2.0
 ; SEQ ID NO 4
 ; LENGTH: 466
 ; TYPE: PRT
 ; ORGANISM: Mus sp.

This Page Blank (uspto)

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: September 30, 2004, 10:58:43 ; Search time 63.1429 Seconds

(without alignments)
149,907 Million cell updates/sec

Title: US-09-988-851A-11

Perfect score: 151
Sequence: 1 DEWVSLAVDAQHAAKRVASGLRFFRLNA 30

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 1017041 seqs, 315518202 residues

Total number of hits satisfying chosen parameters: 1017041

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :
1: sp_archaea:*
2: sp_bacteria:*
3: sp_fungi:*
4: sp_human:*
5: sp_invertebrate:*
6: sp_mammal:*
7: sp_mhc:*
8: sp_organelle:*
9: sp_phage:*
10: sp_plant:*
11: sp_rodent:*
12: sp_virus:*
13: sp_vertebrate:*
14: sp_unclassified:*
15: sp_virus:*
16: sp_bacteriap:*
17: sp_archaeap:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	148	98.0	873	12	Q68286 human cytom
2	121	80.1	902	12	Q68R29 chimpanzee
3	93	61.6	724	12	Q7EFK1 rheus cyto
4	56	37.1	428	2	Q9KH61 mycobacteri
5	55	36.4	1998	12	Q56050 rhopalosiph
6	53	35.1	1318	9	Q8W5T9 bacterioph
7	53	35.1	1318	9	Q8W5T9 Yersinia pe
8	52	34.4	311	2	Q52426 pseudomonas
9	52	34.4	311	2	Q52426 pseudomonas
10	52	34.4	311	2	Q52426 pseudomonas
11	52	34.4	429	16	Q7TX32 mycobacteri
12	52	34.4	454	16	Q53329 mycobacteri
13	52	34.4	818	2	Q7WVY0 alcaligenes
14	51	33.8	205	16	Q87YV9 pseudomonas
15	51	33.8	219	4	Q9Y656 homo sapien
16	51	33.8	269	16	Q7TV91 prochloroc

17	51	33.8	320	12	Q6QND3	Q6QND3 ectocarpus
18	51	33.8	421	4	Q8NC76	Q8NC76 homo sapien
19	50.5	33.4	497	2	Q93PA1	Q93PA1 microscilla
20	50.5	33.4	757	10	Q9ZUJ3	Q9ZUJ3 arbidopsis
21	50	33.1	387	5	Q9GRJ4	Q9GRJ4 leishmania
22	50	33.1	421	13	Q9W7K9	Q9W7K9 gallus gall
23	50	33.1	481	3	Q873C7	Q873C7 neospora
24	50	33.1	670	5	Q95PT0	Q95PT0 leishmania
25	49	32.5	348	16	Q8XB08	Q8XB08 chlorobium
26	49	32.5	407	10	Q9ZOA3	Q9ZOA3 arbidopsis
27	49	32.5	407	10	Q84W03	Q84W03 streptomyces
28	49	32.5	455	2	Q81IC3	Q81IC3 pasteurella
29	49	32.5	461	16	Q8XW22	Q8XW22 clostridium
30	49	32.5	629	16	Q7W0H4	Q7W0H4 bordetella
31	49	32.5	634	16	Q7WC93	Q7WC93 bordetella
32	49	32.5	3498	5	Q20497	Q20497 caenorhabdit
33	48.5	32.1	501	2	Q7WYP9	Q7WYP9 myxococcus
34	48	31.8	98	16	Q9BAJ7	Q9BAJ7 rhizobium
35	48	31.8	179	11	Q8VB81	Q8VB81 mus musculu
36	48	31.8	179	11	Q8C841	Q8C841 mus musculu
37	48	31.8	195	16	Q8RHT7	Q8RHT7 fusobacteri
38	48	31.8	287	16	Q8R7W0	Q8R7W0 streptococc
39	48	31.8	314	2	Q9FD58	Q9FD58 streptococc
40	48	31.8	374	16	Q9BJ72	Q9BJ72 rhizobium
41	48	31.8	415	16	Q9K428	Q9K428 streptomyces
42	48	31.8	421	11	Q7TMU2	Q7TMU2 mus musculu
43	48	31.8	445	16	Q82N38	Q82N38 streptomyces
44	48	31.8	471	16	Q7WPW6	Q7WPW6 bordetella
45	48	31.8	471	16	Q7WBX0	Q7WBX0 bordetella

ALIGNMENTS

RESULT 1	Q68286	PRELIMINARY;	PRT;	873 AA.
ID	Q68286;			
AC	Q68286;			
DT	01-NOV-1996 (TRENBLREL. 01, Last sequence update)			
DT	01-NOV-1996 (TRENBLREL. 01, Last sequence update)			
DT	01-OCT-2002 (TRENBLREL. 22, Last annotation update)			
DE	UL102 protein.			
OS	Human cytomegalovirus (strain AD169).			
OC	Viruses; dsDNA viruses, no RNA stage; Herpesviridae;			
OC	Betaherpesvirinae; Cytomegalovirus.			
OX	NCBI_TaxID=10360;			
RN	[1]			
RP	SEQUENCE FROM N.A.			
RC	STRAIN=AD169;			
RX	MEDLINE=95156603; PubMed=7853511;			
RA	Smith J.A., Pari G.S.;			
RT	"Human cytomegalovirus UL102 gene.";			
RL	J. Virol. 68:3734-1740(1995).			
DR	EMBL:U18285; AAA67889.1;			
SQ	SEQUENCE 873 AA; 94085 MW; 04D5A43015CB16C4 CRC64;			
Query Match	98.0%;	Score 148;	DB 12;	Length 873;
Best Local Similarity	96.7%;	Pred. No. 1.7e-13;		
Matches	29;	Conservative	1;	Mismatches 0;
		Indels	0;	Gaps 0;
QY	1 DEWVSLAVDAQHAAKRVASGLRFFRLNA 30			
DB	844 DEWVSLAVDAQHAAKRVASGLRFFRLNA 873			
RESULT 2	Q68R29	PRELIMINARY;	PRT;	902 AA.
ID	Q68R29;			
AC	Q68R29;			
DT	01-JUN-2002 (TRENBLREL. 21, Created)			
DT	01-JUN-2002 (TRENBLREL. 21, Last sequence update)			
DT	01-JUN-2003 (TRENBLREL. 24, Last annotation update)			
DE	DNA helicase-primase complex protein UL102.			

OS Chimpanzee cytomegalovirus.
OC Viruses; dsDNA viruses, no RNA stage; Herpesviridae;
OC Betaherpesvirinae; Cytomegalovirus.
OX NCBI_TaxID=188763;

RN [1]
RP SEQUENCE FROM N.A.
RA Davidson A.J., Akter P., Dolan A., Wright K.M., Addison C.,
RA Alencor D.J., Hayward G.S., McGeoch D.J.,
RT "The human cytomegalovirus genome revisited."
RL Submitted (FEB-2002) to the EMBL/GenBank/DBJ databases.
DR EMBL: AF480884; AM00739.1; -
DR GO: GO:0004386; F:helicase activity; IEA.
KW Helicase.
SQ SEQUENCE 902 AA; 96000 MW; 8C4853ADE58612A9 CRC64;

Query Match 80.1%; Score 121; DB 12; Length 902;
Best Local Similarity 80.0%; Pred. No. 2.2e-09;
Matches 24; Conservative 2; Mismatches 4; Indels 0; Gaps 0;

OY 1 DEWTRSLAVDAQHAKRVASGGLFFRLN 30
DB 873 DEWARSIGSDAKARVASEGLFFRLN 902

RESULT 3

ID Q7FFK1 PRELIMINARY; PRT; 724 AA.
AC Q7FFK1;
DT 01-OCT-2003 (TREMBlrel. 25, Created)
DT 01-OCT-2003 (TREMBlrel. 25, Last sequence update)
DE 01-OCT-2003 (TREMBlrel. 25, Last annotation update)
DE Rh13.
OS Rhesus cytomegalovirus (strain 68-1) (RhCMV).
OC Viruses; dsDNA viruses, no RNA stage; Herpesviridae;
OC Betaherpesvirinae; Cytomegalovirus.
OX NCBI_TaxID=103930;
RN [1]
RP SEQUENCE FROM N.A.
RC STPAIN=68-1;
RX PubMed=12767982;
RA Hansen S.G., Strelow L.I., Franchi D.C., Anders D.G., Wong S.W.,
RT "Complete Sequence and Genomic Analysis of Rhesus Cytomegalovirus."
RL J. Virol. 77:6620-6636(2003).
DR EMBL: AY186194; AAP50663.1; -
SQ SEQUENCE 724 AA; 80409 MW; 976CA04553C7E1BC CRC64;

Query Match 61.6%; Score 93; DB 12; Length 724;
Best Local Similarity 58.6%; Pred. No. 3e-05;
Matches 17; Conservative 8; Mismatches 4; Indels 0; Gaps 0;

OY 1 DEWTRSLAVDAQHAKRVASGGLFFRLN 29
DB 695 EEWRLSLGSRQAERVAEGSRFFRLN 723

RESULT 4

ID Q9KH61 PRELIMINARY; PRT; 428 AA.
AC Q9KH61;
DT 01-OCT-2000 (TREMBlrel. 15, Created)
DT 01-OCT-2000 (TREMBlrel. 15, Last sequence update)
DT 01-OCT-2000 (TREMBlrel. 15, Last annotation update)
DE Putative ATP/GTP binding protein.
OS Mycobacterium smegmatis.
OC Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;
OC Corynebacterineae; Mycobacteriaceae; Mycobacterium.
OX NCBI_TaxID=1772;
RN [1]
RP SEQUENCE FROM N.A.
RA Sliker D., Nagata V.,
RT "Regulator of a two component system from Mycobacterium smegmatis."
RL Submitted (MAY-2000) to the EMBL/GenBank/DBJ databases.
DR EMBL: AF271296; AAF6191.1; -

SQ SEQUENCE 428 AA; 47506 MW; A2E39D753C993CCF CRC64;
Query Match 37.1%; Score 56; DB 2; Length 428;
Best Local Similarity 52.4%; Pred. No. 6.8;
Matches 11; Conservative 3; Mismatches 7; Indels 0; Gaps 0;

OY 3 WTRSLAVDAQHAKRVASEGL 23
DB 52 WLRSLDVERASARADEGL 72

RESULT 5

ID O56050 PRELIMINARY; PRT; 1998 AA.
AC O56050;
DT 01-JUN-1998 (TREMBlrel. 06, Created)
DT 01-JUN-1998 (TREMBlrel. 06, Last sequence update)
DT 01-OCT-2003 (TREMBlrel. 25, Last annotation update)
DE Nonstructural polyprotein.
OS Rhopalosiphum padi virus.
OC Viruses; ssRNA positive-strand viruses, no DNA stage; Dicistroviridae;
OC Crivavirina.
OX NCBI_TaxID=66834;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=988201645; PubMed=9527915;
RA Moon J.S., Domier L.L., McCoppin N.K., D'Arcy C.J., Jin H.,
RT "Nucleotide sequence analysis shows that Rhopalosiphum padi virus is a
member of a novel group of insect-infecting RNA viruses."
RL Virology 243:54-65(1998).
RN [2]
RP SEQUENCE FROM N.A.
RA Domier L.L., Moon J.S., McCoppin N.K., Jin H.,
RL Submitted (SEP-1997) to the EMBL/GenBank/DBJ databases.
DR EMBL: AF022937; AAC95509.1; -
DR PIR: T08822; T08822.
DR GO: GO:0005524; F:ATP binding; IEA.
DR GO: GO:0003724; F:RNA helicase activity; IEA.
DR GO: GO:0003968; F:RNA-directed RNA polymerase activity; IEA.
DR GO: GO:0015992; P:proton transport; IEA.
DR GO: GO:0006350; P:transcription; IEA.
DR GO: GO:0019079; P:viral genome replication; IEA.
DR InterPro: IPR001914; ATPase_a/bcentr.
DR InterPro: IPR009003; Cys_Ser_trypsin.
DR InterPro: IPR000605; RNA helicase.
DR InterPro: IPR007095; RNA_pol_DS_PS.
DR InterPro: IPR001205; RNA_pol_P3D.
DR InterPro: IPR007094; RNA_pol_PSVir.
DR Pfam: PF00680; RNA_dep_RNA_pol. 1.
DR Pfam: PF00910; RNA_helicase; 1.
DR PROSITE: PS00152; ATPase_ALPHA_BETA; 1.
KW Polyprotein.
SQ SEQUENCE 1998 AA; 228568 MW; 22D5CBA1CA5483D CRC64;

Query Match 36.4%; Score 55; DB 12; Length 1998;
Best Local Similarity 50.0%; Pred. No. 59;
Matches 10; Conservative 5; Mismatches 5; Indels 0; Gaps 0;

OY 9 VDAQHAKRVASEGLFFRL 28
DB 501 LETQHAGRMVSEGLRLMRI 520

RESULT 6

ID O8M5T9 PRELIMINARY; PRT; 1318 AA.
AC O8M5T9;
DT 01-MAR-2002 (TREMBlrel. 20, Created)
DT 01-MAR-2002 (TREMBlrel. 20, Last sequence update)
DT 01-OCT-2003 (TREMBlrel. 25, Last annotation update)
DE Internal virion protein D.
GN GENE 16.
OS Bacteriophage T3.


```

OC Viruses; dsDNA viruses, no RNA stage; Caudovirales; Podoviridae;
OC T7-like viruses.
OX NCBI_TaxId=10759;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=Luria;
RA Pajunen M.I.;
RN Submitted (JUN-2001) to the EMBL/GenBank/DBJ databases.
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=Luria;
RA Pajunen M.I., Elizondo M.R., Skurnik M., Kielczawa J., Molinex I.J.;
RT "Complete nucleotide sequence and likely recombinatorial origin of
RT bacteriophage T3."
RL J. Mol. Biol. 0:0-0(0).
DR EMBL; AJ318471; CAC86304.1; -.
DR InterPro; IPR008258; SLT dom.
DR InterPro; IPR00189; Transglyc_AS.
PFam; PF01464; SLT; 1.
DR PROSITE; PS00922; TRANSGLYCOSYLASE; 1.
SQ SEQUENCE 1318 AA; 144017 MW; 6533E46B1C818F15 CRC64;

Query Match 35.1%; Score 53; DB 9; Length 1318;
Best Local Similarity 60.0%; Pred. No. 73;
Matches 12; Conservative 3; Mismatches 5; Indels 0; Gaps 0;

CY 5 RSLAVDQHAARKVASEGLR 24
Db 358 KALVGAESAALNVASEGLR 377

RESULT 7
ID 085809 PRELIMINARY; PRT; 1318 AA.
AC 085809;
DT 01-JUN-2003 (TREMBlrel. 24, Created)
DT 01-JUN-2003 (TREMBlrel. 24, Last sequence update)
DT 01-OCT-2003 (TREMBlrel. 25, Last annotation update)
DE Internal virion protein.
OS Versinia pestis phage phiA1122.
OC Viruses; dsDNA viruses, no RNA stage; Caudovirales; Podoviridae;
OC T7-like viruses.
OX NCBI_TaxId=227720;
RN [1]
RP SEQUENCE FROM N.A.
RA Garcia E., Elliott J.M., Ramanculov E., Chain P.S., Chu M.C.,
RA Molinex I.J.;
RT "The genome sequence of Versinia pestis bacteriophage phiA1122 reveals
RT an intimate history with the coliphage T3."
RL Submitted (FEB-2003) to the EMBL/GenBank/DBJ databases.
DR EMBL; AY247822; AAP20539.1; -.
DR InterPro; IPR008258; SLT dom.
DR InterPro; IPR00189; Transglyc_AS.
PFam; PF01464; SLT; 1.
DR PROSITE; PS00922; TRANSGLYCOSYLASE; 1.
SQ SEQUENCE 1318 AA; 144004 MW; 084C0CDEB623A9DE CRC64;

Query Match 35.1%; Score 53; DB 9; Length 1318;
Best Local Similarity 60.0%; Pred. No. 73;
Matches 12; Conservative 3; Mismatches 5; Indels 0; Gaps 0;

CY 5 RSLAVDQHAARKVASEGLR 24
Db 358 KALVGAESAALNVASEGLR 377

RESULT 8
ID 052426 PRELIMINARY; PRT; 311 AA.
AC 052426;
DT 01-NOV-1996 (TREMBlrel. 01, Created)
DT 01-NOV-1996 (TREMBlrel. 01, Last sequence update)
DT 01-OCT-2003 (TREMBlrel. 25, Last annotation update)

```

```

DE Avirulence gene D (Fragment).
GN AVR.
OS Pseudomonas syringae.
OC Bacteria; Proteobacteria; Gammaproteobacteria; Pseudomonadales;
OC Pseudomonadaceae; Pseudomonas.
OX NCBI_TaxId=317;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=94220735; PubMed=8167364;
RA Yucel I., Boyd C., Debnam Q., Keen N.T.;
RT "Two different classes of avr alleles occur in pathovars of
RT Pseudomonas syringae."
RL Mol. Plant Microbe Interact. 7:131-139 (1994).
DR EMBL; L11334; AAA20577.2; -.
DR InterPro; IPR008799; Pseudomon_AVRD.
DR Pfam; PF05655; Pseudomon_AVRD; 1.
FT NON TER 311
SQ SEQUENCE 311 AA; 34465 MW; 9CFB08B50C05378B CRC64;

Query Match 34.4%; Score 52; DB 2; Length 311;
Best Local Similarity 47.6%; Pred. No. 19;
Matches 10; Conservative 6; Mismatches 5; Indels 0; Gaps 0;

CY 3 WYRSLAVDQHAARKVASEGL 23
Db 249 WYRSLNITIAENPAKRIATRL 269

RESULT 9
ID 086072 PRELIMINARY; PRT; 311 AA.
AC 086072;
DT 01-NOV-1998 (TREMBlrel. 08, Created)
DT 01-NOV-1998 (TREMBlrel. 08, Last sequence update)
DT 01-OCT-2003 (TREMBlrel. 25, Last annotation update)
DE Avirulence D protein.
GN AVR.
OS Pseudomonas syringae pv. apti.
OC Bacteria; Proteobacteria; Gammaproteobacteria; Pseudomonadales;
OC Pseudomonadaceae; Pseudomonas.
OX NCBI_TaxId=81036;
RN [1]
RP SEQUENCE FROM N.A.
RA Keen N.T., Boyd C.M.;
RL Submitted (AUG-1998) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF083919; AAC3120.1; -.
DR InterPro; IPR008799; Pseudomon_AVRD.
DR Pfam; PF05655; Pseudomon_AVRD; 1.
SQ SEQUENCE 311 AA; 34768 MW; 11783BD4642D7F5 CRC64;

Query Match 34.4%; Score 52; DB 2; Length 311;
Best Local Similarity 47.6%; Pred. No. 19;
Matches 10; Conservative 6; Mismatches 5; Indels 0; Gaps 0;

CY 3 WYRSLAVDQHAARKVASEGL 23
Db 249 WYRSLNITIAENPAKRIATRL 269

RESULT 10
ID 086071 PRELIMINARY; PRT; 311 AA.
AC 086071;
DT 01-NOV-1998 (TREMBlrel. 08, Created)
DT 01-NOV-1998 (TREMBlrel. 08, Last sequence update)
DT 01-OCT-2003 (TREMBlrel. 25, Last annotation update)
DE Avirulence D protein.
GN AVR.
OS Pseudomonas syringae pv. cilarro.
OC Bacteria; Proteobacteria; Gammaproteobacteria; Pseudomonadales;
OC Pseudomonadaceae; Pseudomonas.
OX NCBI_TaxId=81035;
RN [1]

```

RP SEQUENCE FROM N.A.
 RA Keen N.T., Boyd C.M.;
 RL Submitted (Aug-1998) to the EMBL/GenBank/DBJ databases.
 DR EMBL: AF083918; AAC3119.1; -;
 DR InterPro: IPR006799; Pseudomon_Ayrd.
 DR Pfam: PF05655; Pseudomon_Ayrd_1
 SQ SEQUENCE 311 AA; 34546 MW; E1639917A95EC6A0 CRC64;

Query Match 34.4%; Score 52; DB 2; Length 311;
 Best Local Similarity 47.6%; Pred. No. 19;
 Matches 10; Conservative 6; Mismatches 5; Indels 0; Gaps 0;

Qy 3 WWSLAVDAQHAKRVASEGL 23
 Db 249 WWSLNTIAENPAKRIATRL 269

RESULT 11

ID Q7TX32 PRELIMINARY; PRT; 429 AA.
 AC Q7TX32;
 DT 01-OCT-2003 (TREMblrel. 25, Created)
 DT 01-OCT-2003 (TREMblrel. 25, Last sequence update)
 DT 01-OCT-2003 (TREMblrel. 25, Last annotation update)
 DE Conserved hypothetical protein.
 GN MB3205
 OS Mycobacterium bovis.
 OC Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;
 OC Corynebacterineae; Mycobacteriaceae; Mycobacterium;
 OX NCBI_TaxID=1765;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=AF2122/97;
 RX MEDLINE=22709107; PubMed=12768972;
 RA Garnier T., Elgimeleer K., Camus J.-C., Medina N., Mansoor H.,
 RA Pryor M., Duthoy S., Gordin S., Lacroix C., Monsemp C., Simon S.,
 RA Harris B., Atkin R., Doggett J., Mayes R., Keating L., Wheeler P.R.,
 RA Parkhill J., Barrall B.G., Cole S.T., Gordon S.V., Hewinson R.G.;
 RT "The complete genome sequence of Mycobacterium bovis."
 RL Proc. Natl. Acad. Sci. U.S.A. 100:7677-7682(2003).
 KW EMBL: BX248345; CAD95297.1; -;
 KM Complete proteome.
 SQ SEQUENCE 429 AA; 48053 MW; B662D9374654987C CRC64;

Query Match 34.4%; Score 52; DB 16; Length 429;
 Best Local Similarity 50.0%; Pred. No. 27;
 Matches 10; Conservative 3; Mismatches 7; Indels 0; Gaps 0;

Qy 3 WWSLAVDAQHAKRVASEG 22
 Db 53 WWSLVDVVERASARADPEG 72

RESULT 12

ID Q53329 PRELIMINARY; PRT; 454 AA.
 AC Q53329;
 DT 01-JUN-1998 (TREMblrel. 06, Created)
 DT 01-MAR-2002 (TREMblrel. 20, Last sequence update)
 DT 01-OCT-2003 (TREMblrel. 25, Last annotation update)
 DE Hypothetical protein RV3179.
 GN RV3179 OR MT3270 OR MYO14.23.
 OS Mycobacterium tuberculosis.
 OC Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;
 OC Corynebacterineae; Mycobacteriaceae; Mycobacterium;
 OX NCBI_TaxID=1773;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=H3TRV;
 RX MEDLINE=98295987; PubMed=9634230;
 RA Cole S.T., Brosch R., Parkhill J., Garnier T., Churcher C., Harris D.,
 RA Gordon S.V., Eigemeier K., Gas S., Barry C.E. III, Tekala F.,
 RA Badcock K., Basham D., Brown D., Chillingworth T., Connor R.,

RA Davies R., Devlin K., Feltwell T., Gentles S., Hamlin N., Holroyd S.,
 RA Hornsby T., Jagers K., Krogh A., McLean J., Moule S., Murphy L.,
 RA Oliver S., Osborne K., Quail M.A., Rajandream M.A., Rogers J.,
 RA Rutter S., Seeger K., Skelton S., Squares S., Squares R.,
 RA Sulten J.E., Taylor K., Whitehead S., Barrall B.G.;
 RT "Deciphering the biology of Mycobacterium tuberculosis from the
 complete genome sequence."
 RL Nature 393:537-544(1998).
 RN [2]

RP SEQUENCE FROM N.A.
 RC STRAIN=CDC 1551 / Oshkosh;
 RA Fleischnann R.D., Alland D., Eisen J.A., Carpenter L., White O.,
 RA Peterson J., Debay R., Dodson R., Gwinn M., Haft D., Hickey E.,
 RA Kolonay J.F., Nelson M.C., Umayan L.A., Ermolaeva M., Salzberg S.L.,
 RA Delcher A., Utterback T., Weidman J., Kouri H., Gill J., Mikula A.,
 RA Bisht W.;
 RT "Whole genome comparison of Mycobacterium tuberculosis clinical and
 laboratory strains."
 RL Submitted (APR-2001) to the EMBL/GenBank/DBJ databases.
 DR EMBL: AL021646; CA16644.1; ALT_INIT.
 DR EMBL: AB007140; AAK47609.1; -;
 DR FIR: A70949; A70949.
 DR TIGR: MT3270; -;
 DR Tuberculist; RV3179;
 KW Hypothetical protein; Complete proteome.
 SQ SEQUENCE 454 AA; 50934 MW; 7C63015FD328ED0 CRC64;

Query Match 34.4%; Score 52; DB 16; Length 454;
 Best Local Similarity 50.0%; Pred. No. 29;
 Matches 10; Conservative 3; Mismatches 7; Indels 0; Gaps 0;

Qy 3 WWSLAVDAQHAKRVASEG 22
 Db 76 WWSLVDVVERASARADPEG 97

RESULT 13

ID Q7WVY0 PRELIMINARY; PRT; 818 AA.
 AC Q7WVY0;
 DT 01-OCT-2003 (TREMblrel. 25, Created)
 DT 01-OCT-2003 (TREMblrel. 25, Last sequence update)
 DT 01-OCT-2003 (TREMblrel. 25, Last annotation update)
 DE Putative tRb-Like protein.
 GN TRBE OR PHG362.
 OS Alcaligenes eutrophus (Ralstonia eutropha).
 OG Plasmid megaplasmid PHG1.
 OC Bacteria; Proteobacteria; Betaproteobacteria; Burkholderiales;
 OC Burkholderiaceae; Ralstonia.
 OX NCBI_TaxID=510;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=H16;
 RA Schwartz E., Henne A., Cramm R., Eitinger T., Friedrich B.,
 RA Gottschalk G.;
 RT "Complete Nucleotide Sequence of pHG1: A Ralstonia eutropha H16
 RT Megaplasmid Encoding Key Enzymes of H2-based Lithoautotrophy and
 RT Anaerobiosis."
 RL Submitted (May-2003) to the EMBL/GenBank/DBJ databases.
 DR EMBL: AY305378; AAP86111.1; -;
 KW Plasmid.

SQ SEQUENCE 818 AA; 92392 MW; AFD46F7661EDC99B CRC64;

Query Match 34.4%; Score 52; DB 2; Length 818;
 Best Local Similarity 45.8%; Pred. No. 59;
 Matches 11; Conservative 5; Mismatches 8; Indels 0; Gaps 0;

Qy 1 DEWWSLAVDAQHAKRVASEGL 24
 Db 324 DEGREELAKGKHALKLTAGVQ 347

RESULT 14

This Page Blank (uspto)

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: September 30, 2004, 10:58:43 ; Search time 54.7238 Seconds
(without alignments)
149.907 Million cell updates/sec

Title: US-09-988-851A-12
Perfect score: 135
Sequence: 1 TWLEERDEWVRSILAVDAQHARVAVS 26

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 1017041 seqs, 315518202 residues

Total number of hits satisfying chosen parameters: 1017041

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : SPTREMBL_25:*
1: sp_archaea:*
2: sp_bacteria:*
3: sp_fungi:*
4: sp_human:*
5: sp_invertebrate:*
6: sp_mammal:*
7: sp_mhc:*
8: sp_organelle:*
9: sp_phage:*
10: sp_plant:*
11: sp_rodent:*
12: sp_virus:*
13: sp_vertebrate:*
14: sp_unclassified:*
15: sp_virus:*
16: sp_bacteriap:*
17: sp_archaeap:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	135	100.0	873	12	Q68286 human cytom
2	88	65.2	902	12	Q8QR29 chimpanzee
3	70	51.9	724	12	Q7TFK1 thesus cyto
4	55	40.7	477	2	Q9FAX3 flexibacter
5	55	40.7	477	2	Q9FAU8 flexibacter
6	55	40.7	1078	5	Q9N6M7 leishmania
7	54	40.0	129	16	Q8ZRN6 salmonella
8	54	40.0	129	16	Q8ZRN6 salmonella
9	53.5	39.6	803	3	Q9P682 neurospora
10	53	39.3	390	2	Q8RQX3 marine cfb
11	53	39.3	478	2	Q9AJL2 marine cfb
12	53	39.3	621	16	Q7UL57 rhodospirillum rubrum
13	52	38.5	311	2	Q86072 pseudomonas
14	51	37.8	333	13	Q9DE09 gallus galli
15	51	37.8	377	2	Q9RC92 bacillus sp
16	51	37.8	415	6	Q86659 bos taurus

17	51	37.8	2265	5	Q8IID2	Q8IID2 plasmid
18	50.5	37.4	548	2	Q51340	Q51340 pseudomonas
19	50.5	37.4	548	16	Q9I183	Q9I183 pseudomonas
20	50	37.0	203	16	Q9P9P8	Q9P9P8 xylophaga
21	50	37.0	331	16	Q9I192	Q9I192 pseudomonas
22	50	37.0	353	13	Q9W6Q0	Q9W6Q0 gallus galli
23	50	37.0	384	16	Q8XAC2	Q8XAC2 escherichia
24	50	37.0	451	2	Q05074	Q05074 streptomyces
25	49.5	36.7	547	10	Q8L414	Q8L414 xylophaga
26	49	36.3	203	16	Q87CR8	Q87CR8 xylophaga
27	49	36.3	305	2	Q52740	Q52740 corynebacter
28	49	36.3	311	2	Q52426	Q52426 pseudomonas
29	49	36.3	311	2	Q86071	Q86071 pseudomonas
30	49	36.3	347	16	Q9AAX0	Q9AAX0 caulobacter
31	49	36.3	442	9	Q854W2	Q854W2 mycobacteri
32	49	36.3	524	2	Q7X127	Q7X127 xanthomonas
33	49	36.3	524	16	Q8PRK2	Q8PRK2 xanthomonas
34	49	36.3	608	16	Q912B9	Q912B9 pseudomonas
35	49	36.3	675	16	Q82J52	Q82J52 streptomyces
36	49	36.3	1145	3	Q9UR03	Q9UR03 candida alb
37	48.5	35.9	302	16	Q7U323	Q7U323 helicobacte
38	48.5	35.9	526	5	Q9V708	Q9V708 dirosophila
39	48.5	35.9	903	5	Q9XX09	Q9XX09 caenorhabdi
40	48.5	35.9	951	5	Q9U2M2	Q9U2M2 caenorhabdi
41	48.5	35.9	952	5	Q9XZQ3	Q9XZQ3 caenorhabdi
42	48.5	35.9	1107	5	Q9XX14	Q9XX14 caenorhabdi
43	48.5	35.9	1107	5	Q9XZQ4	Q9XZQ4 caenorhabdi
44	48.5	35.9	1746	5	Q22830	Q22830 caenorhabdi
45	48.5	35.6	98	16	Q98A07	Q98A07 rhizobium 1

ALIGNMENTS

RESULT 1

Q68286 ID Q68286 PRELIMINARY; PRT; 873 AA.
AC Q68286;
DT 01-NOV-1996 (TREMBLrel. 01, Created)
DT 01-NOV-1996 (TREMBLrel. 01, Last sequence update)
DT 01-OCT-2002 (TREMBLrel. 22, Last annotation update)
DE U102 protein.
OS Human cytomegalovirus (strain AD169).
OC Viruses; dsDNA viruses, no RNA stage; Herpesviridae;
OC Betaherpesvirinae; Cytomegalovirus.
OX NCBI_TaxID=10360;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=AD169;
RX MEDLINE=95156603; PubMed=7853511;
RA Smith U.A., Part G.S.;
RT "Human cytomegalovirus U102 gene."
RL J. Virol. 69:1734-1740 (1995).
DR EMBL; U18289; AAA67889.1; -
SQ SEQUENCE 873 AA; 94085 MW; 04D5A4301SCB16C4 CRC64;

Query Match 100.0%; Score 135; DB 12; Length 873;
Best Local Similarity 100.0%; Pred. No. 1.1e-10;
Matches 26; Conservative 0; Mismatches 0; Indels 0;
Gaps 0;
Qy 1 TWLEERDEWVRSILAVDAQHARVAVS 26
Db 838 TWLEERDEWVRSILAVDAQHARVAVS 963
RESULT 2
Q68286 ID Q68286 PRELIMINARY; PRT; 902 AA.
AC Q68286;
DT 01-JUN-2002 (TREMBLrel. 21, Created)
DT 01-JUN-2002 (TREMBLrel. 21, Last sequence update)
DT 01-JUN-2003 (TREMBLrel. 24, Last annotation update)
DE DNA helicase-primase complex protein U102.

DR Pfam: PF00204; DNA gyrase3; 1.
 DR Pfam: PF02518; HATPase C; 1.
 DR Pfam: PF01751; Toprim; 1.
 DR PRINTS; PR00418; TP12FAMILY.
 DR SMART; SM00433; TOP2c; 1.
 DR PROSITE; PS00013; PROKAR LIPOPROTEIN; 1.
 DR PROSITE; PS00177; TOPOISOMERASE II; 1.
 DR ATP-binding; Isomerase; Topoisomerase.
 FT NON_TER 1
 FT NON_TER 477
 SQ SEQUENCE 477 AA; 52984 MW; 31F756D1A6299FD4 CRC64;
 Query Match 40.7%; Score 55; DB 2; Length 477;
 Best Local Similarity 44.8%; Pred. No. 16;
 Matches 13; Conservative 3; Mismatches 7; Indels 6; Gaps 1;

QY 1 TWLEERDEWVRS-----LAVDAQHARR 23
 DB 260 TWLEERDEWVRSQIVAKYITLAAQAHARR 288

RESULT 6
 Q9N6M7 PRELIMINARY; PRT; 1078 AA.
 AC Q9N6M7;
 DT 01-OCT-2000 (TREMBLrel. 15, Created)
 DT 01-OCT-2000 (TREMBLrel. 15, Last sequence update)
 DT 01-OCT-2002 (TREMBLrel. 22, Last annotation update)
 DE L712.4.
 GN Leishmania major.
 OS Eukaryote; Euglenozoa; Kinetoplastida; Trypanosomatidae; Leishmania.
 OC Eukaryote; Euglenozoa; Kinetoplastida; Trypanosomatidae; Leishmania.
 OX NCBI_TaxID=5664;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=Fr1edlin;
 RA Worthey E.A., Sisk E., Hixson G., Kiser P., Rickel E., Hassebrock M.,
 RA Cantina J., Sunkin S., Stuart K.D., Wylter P.J.;
 RT "Direct Submission."
 RL Submitted (JUN-2002) to the EMBL/GenBank/DBJ databases.
 DR EMBL; AC125735; AAM69046.1; -
 SQ SEQUENCE 1078 AA; 113007 MW; 326C10F816EFED47 CRC64;
 Query Match 40.7%; Score 55; DB 5; Length 1078;
 Best Local Similarity 36.4%; Pred. No. 39;
 Matches 8; Conservative 7; Mismatches 7; Indels 0; Gaps 0;

QY 2 WLEERDEWVRSIAYDAQHARR 23
 DB 546 WMTQRMWMLQEGODAHASR 567

RESULT 7
 Q8ZRN6 PRELIMINARY; PRT; 129 AA.
 AC Q8ZRN6;
 DT 01-MAR-2002 (TREMBLrel. 20, Created)
 DT 01-MAR-2002 (TREMBLrel. 20, Last sequence update)
 DT 01-JUN-2003 (TREMBLrel. 24, Last annotation update)
 DE Putative laccoylglycathione lyase.
 GN YABR OR STM0235.
 OS Salmonella typhimurium.
 OC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
 OC Enterobacteriaceae; Salmonella.
 OX NCBI_TaxID=602;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=LT2 / SGSC1412 / ATCC 700720;
 RX MEDLINE=21534948; PubMed=11677609;
 RA McClelland M., Sanderson K.E., Spieth J., Clifton S.W., Latreille P.,
 RA Courtney L., Portolillo S., Ali J., Dante M., Du F., Hou S., Layman D.,
 RA Leonard S., Nguyen C., Scott K., Holmes A., Grewal N., Mulvaney E.,
 RA Ryan E., Sun H., Florea L., Miller W., Stoneking T., Nhan W.,

RA Waterston R., Wilson R.K.;
 RT "Complete genome sequence of Salmonella enterica serovar Typhimurium
 LT2."
 RL Nature 413:852-856(2001).
 DR EMBL; AE008705; AAL1919.1; -
 DR GO; GO:0016829; F:lyase activity; IEA.
 DR InterPro; IPR004360; Gly_bleo_diox.
 DR Pfam; PF00903; Glyoxalase; 1.
 KW Lyase; Hypothetical protein; Complete proteome.
 SQ SEQUENCE 129 AA; 14809 MW; FB791AF6AF654902 CRC64;
 Query Match 40.0%; Score 54; DB 16; Length 129;
 Best Local Similarity 50.0%; Pred. No. 5.1;
 Matches 9; Conservative 3; Mismatches 6; Indels 0; Gaps 0;

QY 2 WLEERDEWVRSIAYDAQH 19
 DB 37 WLEERDSWKGDALINGQY 54

RESULT 8
 Q8Z997 PRELIMINARY; PRT; 129 AA.
 AC Q8Z997;
 DT 01-MAR-2002 (TREMBLrel. 20, Created)
 DT 01-MAR-2002 (TREMBLrel. 20, Last sequence update)
 DT 01-JUN-2003 (TREMBLrel. 24, Last annotation update)
 DE Hypothetical protein STY0260.
 GN STY0260 OR T0237.
 GN STY0260 OR T0237.
 OS Salmonella typhi.
 OC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
 OC Enterobacteriaceae; Salmonella.
 OX NCBI_TaxID=601;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=CT18;
 RX MEDLINE=21534947; PubMed=11677608;
 RA Parkhill J., Dougan G., James K.D., Thomson N.R., Pickard D., Main J.,
 RA Churcher C., Mungall K.L., Bentley S.D., Holden M.T.G., Sebatina M.,
 RA Baker S., Basham D., Brooks K., Chillingworth T., Connor P.,
 RA Cronin A., Davis P., Davies R.M., Dowd L., White N., Farrar J.,
 RA Felwell T., Hamlin N., Haque A., Hien T.T., Holroyd S., Jagels K.,
 RA Krogh A., Larsen T.S., Leather S., Moule S., O'Gea P., Parry C.,
 RA Quail M., Rutherford K., Simmonds M., Skelton J., Stevens K.,
 RA Whitehead S., Barrett B.G.;
 RT "Complete genome sequence of a multiple drug resistant Salmonella
 RT enterica serovar Typhi CT18."
 RL Nature 413:848-852(2001).
 RN [2]
 RP SEQUENCE FROM N.A.
 RC STRAIN=TY2 / ATCC 700931;
 RX MEDLINE=22531367; PubMed=12644504;
 RA Deng W., Liou S.-R., Plunkett G. III, Mayhew G.F., Rose D.J.,
 RA Burdand V., Kodoyanni V., Schwartz D.C., Blatner F.R.;
 RT "Comparative genomics of Salmonella enterica serovar Typhi strains Ty2
 RT and CT18."
 RL J. Bacteriol. 185:2330-2337(2003).
 DR EMBL; AF627266; CND08694.1; -
 DR EMBL; AF016834; AAO67966.1; -
 DR InterPro; IPR004360; Gly_bleo_diox.
 DR Pfam; PF00903; Glyoxalase; 1.
 KW Hypothetical protein; Complete proteome.
 SQ SEQUENCE 129 AA; 14781 MW; BB971408BF732E14 CRC64;

Query Match 40.0%; Score 54; DB 16; Length 129;
 Best Local Similarity 50.0%; Pred. No. 5.1;
 Matches 9; Conservative 3; Mismatches 6; Indels 0; Gaps 0;

QY 2 WLEERDEWVRSIAYDAQH 19
 DB 37 WLEERDSWKGDALINGQY 54


```

GN RB9722.
OS Rhodospirillum rubrum.
OC Bacteria; Planctomycetes; Planctomycetacia; Planctomycetales;
OC Planctomycetaceae; Pirellula.
OX NCBI_TaxID=117;
RN
RP SEQUENCE FROM N.A.
RC STRAIN=1;
RC MEDLINE=22735913; PubMed=12835416;
RA Gloeckner F.O., Kube M., Bauer M., Teeling H., Lombardot T.,
RA Ludwig W., Gade D., Beck A., Borzym K., Heilmann K., Rabus R.,
RA Schleutner H., Mann R., Reinhardt R.;
RT "Complete genome sequence of the marine planctomycete Pirellula sp.
RT strain 1."
RL Proc. Natl. Acad. Sci. U.S.A. 100:8298-8303 (2003).
KM EMBL: BX294150; CAD76421.1; -.
KM Hypothetical protein; Complete proteome.
SQ SEQUENCE 621 AA; 67832 MW; 0C92C807C14920C9 CRC64;

Query Match 39.3%; Score 53; DB 16; Length 621;
Best Local Similarity 48.0%; Pred. No. 41;
Matches 12; Conservative 6; Mismatches 5; Indels 2; Gaps 1;

QY 1 TWLEERDEWVSLAVDAQHARRVA 25
DB 100 TWVGERIEATQNL--NRQHAIRLVA 122

RESULT 13
086072 ID 086072 PRELIMINARY; PRT; 311 AA.
AC 086072;
DT 01-NOV-1998 (TREMBlrel. 08, Created)
DT 01-NOV-1998 (TREMBlrel. 08, Last sequence update)
DT 01-OCT-2003 (TREMBlrel. 25, Last annotation update)
DE Avirulence D protein.
GN AVR-D.
OS Pseudomonas syringae pv. apii.
OC Bacteria; Proteobacteria; Gammaproteobacteria; Pseudomonadales;
OC Pseudomonadaceae; Pseudomonas.
OX NCBI_TaxID=81036;
RN [1]
RP SEQUENCE FROM N.A.
RP Keen N.T., Boyd C.M.;
RA Submitted (AUG-1998) to the EMBL/GenBank/DBJ databases.
RA EMBL: AF083919; AAC33120.1; -.
DR InterPro: IPR006799; Pseudomon_AVRD.
DR Pfam: PF05655; Pseudomon_AVRD.1.
SQ SEQUENCE 311 AA; 34768 MW; 1178B3ED4642D7F5 CRC64;

Query Match 38.5%; Score 52; DB 2; Length 311;
Best Local Similarity 44.0%; Pred. No. 26;
Matches 11; Conservative 7; Mismatches 3; Indels 4; Gaps 1;

QY 6 RDE---WVSLAVDAQHARRVA 26
DB 242 RDESNLTLMRSINTIAENPAKRIA 266

RESULT 14
Q9DE09 ID Q9DE09 PRELIMINARY; PRT; 333 AA.
AC Q9DE09;
DT 01-MAR-2001 (TREMBlrel. 16, Created)
DT 01-MAR-2001 (TREMBlrel. 16, Last sequence update)
DT 01-JUN-2003 (TREMBlrel. 24, Last annotation update)
DE Homeodomain protein GH6.
OS Gallus gallus (Chicken).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Archosauromia; Aves; Neognathae; Galliformes; Phasianinae;
OC Gallus.
OX NCBI_TaxID=9031;
RN [1]

```

```

RP SEQUENCE FROM N.A.
RX MEDLINE=20515605; PubMed=11060230;
RA Schulte D., Cepko C.L.;
RT "Two homeobox genes define the domain of EphA3 expression in the
RT developing chick retina."
RT development 127:5033-5045 (2000).
CC -1- SUBCELLULAR LOCATION: NUCLEAR (BY SIMILARITY).
DR EMBL: AF227921; AAG46561.1; -.
DR HSSP: P23441; 1FTT.
DR GO: GO:0005634; C:nucleus; IEA.
DR GO: GO:0003700; F:transcription factor activity; IEA.
DR GO: GO:0006355; P:regulation of transcription, DNA-dependent; IEA.
DR InterPro: IPR001356; Homeobox.
DR InterPro: IPR000047; HTH_lambdarepressr.
DR Pfam: PF00046; homeobox_1.
DR PRINTS: PR00024; HOMEBOX.
DR PRINTS: PR00031; HTHREPRESSR.
DR ProDom: PD000010; Homeobox; 1.
DR SMART: SM00389; HOX; 1.
DR PROSITE: PS00027; HOMEBOX_1; 1.
DR PROSITE: PS00071; HOMEBOX_2; 1.
DR DNA-binding; Homeobox; Nuclear protein.
SQ SEQUENCE 333 AA; 34896 MW; 4B0EB64904022B12 CRC64;

Query Match 37.8%; Score 51; DB 13; Length 333;
Best Local Similarity 37.9%; Pred. No. 40;
Matches 11; Conservative 6; Mismatches 6; Indels 6; Gaps 2;

QY 2 WLE-ERDEWVSLAVDAQ----HARRV 24
DB 248 WFONRNRKWKQLADLEAMLSHAQR 276

RESULT 15
Q9RC92 ID Q9RC92 PRELIMINARY; PRT; 377 AA.
AC Q9RC92;
DT 01-MAY-2000 (TREMBlrel. 13, Created)
DT 01-MAY-2000 (TREMBlrel. 13, Last sequence update)
DT 01-JUN-2003 (TREMBlrel. 24, Last annotation update)
DE Unsaturated glucuronyl hydrolase.
GN UG.
OS Bacillus sp. G11.
OC Bacteria; Firmicutes; Bacillales; Bacillaceae; Bacillus.
OX NCBI_TaxID=84635;
RN [1]
RP SEQUENCE FROM N.A.
RP STRAIN=G11;
RC MEDLINE=99373108; PubMed=10441389;
RA Hashimoto W., Kobayashi E., Nankai H., Sato N., Miya T., Kawai S.,
RA Murata K.;
RT "Unsaturated glucuronyl hydrolase of Bacillus sp. G11: Novel enzyme
RT prerequisite for metabolism of unsaturated oligosaccharides produced
RT by polysaccharide lyases."
RL Arch. Biochem. Biophys. 368:367-374 (1999).
RL EMBL: AB019619; BA84216.1; -.
DR GO: GO:0016787; F:hydrolase activity; IEA.
DR Hydrolyase.
SQ SEQUENCE 377 AA; 42861 MW; 430593B89216680 CRC64;

Query Match 37.8%; Score 51; DB 2; Length 377;
Best Local Similarity 45.5%; Pred. No. 46;
Matches 10; Conservative 3; Mismatches 9; Indels 0; Gaps 0;

QY 2 WLEERDEWVSLAVDAQHARR 23
DB 101 WYKDESAKRLADADAVLMR 122

Search completed: September 30, 2004, 11:08:39
Job time : 57.7238 secs

```

This Page Blank (uspto)

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: September 30, 2004, 10:58:43 ; Search time 9.65714 Seconds
(without alignments)
140.189 Million cell updates/sec

Title: US-09-988-851A-12

Perfect score: 135
Sequence: 1 TWLEERDEVRSIADVAQHARRVAS 26

Scoring table: BLOSUM62
Gapop 10.0, Gapext 0.5

Searched: 141681 seqs, 52070155 residues
Total number of hits satisfying chosen parameters: 141681

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database: SwissProt_42:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	132	97.8	798	HEPA_HCMVA	P16827 human cytom
2	51	37.8	428	GFAP_BOVIN	Q28115 bos taurus
3	51	37.8	2470	TORI_YEAST	P31519 saccharomyc
4	51	37.8	2473	TORI_YEAST	P31519 saccharomyc
5	50.5	37.4	861	OPGH_PSEAE	Q9hu66 pseudomonas
6	49	36.3	940	SPR1_MYCLE	P57996 mycobacteri
7	49	36.3	2388	SPCP_RAT	Q9gw28 rattus norv
8	49	36.3	2390	SPCP_HUMAN	O15020 homo sapien
9	48	35.6	719	P5CS_MESCR	O65361 m delia i-p
10	47.5	35.2	364	FAH2_VIBRA	O87h12 vibrio para
11	47	34.8	350	PLEK_HUMAN	P08567 homo sapien
12	47	34.8	350	PLEK_MOUSE	O9jhk5 mus musculu
13	47	34.8	448	VIME_CRICK	P08670 mus musculu
14	47	34.8	459	VIME_CRICK	P08670 mus musculu
15	47	34.8	464	VIME_MESAU	P22544 mesocricetu
16	47	34.8	465	VIME_BOVIN	P48616 bos taurus
17	47	34.8	465	VIME_HUMAN	P08670 homo sapien
18	47	34.8	465	VIME_MOUSE	P01052 mus musculu
19	47	34.8	465	VIME_RAT	P10100 rattus norv
20	46	34.1	430	GFAP_MOUSE	P03995 mus musculu
21	46	34.1	450	VIME_CARAU	P48673 carassius a
22	46	34.1	455	VIME_CYPCA	Q22155 cyprinus a
23	46	34.1	468	PERI_RAT	P41807 rattus norv
24	46	34.1	471	PERI_HUMAN	P41807 rattus norv
25	46	34.1	475	PERI_MOUSE	P15331 mus musculu
26	46	34.1	508	GOBI_MYCTU	O10677 mycobacteri
27	46	34.1	1556	GFAP_MOUSE	O84498 diosophilla
28	45.5	33.7	400	CYH2_HUMAN	Q99418 homo sapien
29	45.5	33.7	400	CYH2_MOUSE	P97693 mus musculu
30	45.5	33.7	4427	PRSL_BACSV	Q05470 bacillus su
31	45	33.3	359	GFAP_CARAU	P48673 carassius a
32	45	33.3	481	AKT2_RAT	P47197 rattus norv
33	45	33.3	554	YJUK_ECOLI	P37797 escherichia

ALIGNMENTS

RESULT 1	ID	HEPA_HCMVA	STANDARD;	PRT;	798 AA.
34	45	33.3	567	1	ARAB_VIBRA
35	45	33.3	1117	1	YN96_YEAST
36	44.5	33.0	536	1	OPGD_XYLPT
37	44.5	33.0	650	1	ACSA_SHEON
38	44	32.6	241	1	HIS4_BIFLO
39	44	32.6	352	1	PSBD_CHURE
40	44	32.6	376	1	O43A_DROME
41	44	32.6	394	1	CYH4_HUMAN
42	44	32.6	432	1	GFAP_HUMAN
43	44	32.6	586	1	YNG5_YEAST
44	44	32.6	649	1	RA32_SCHPO
45	44	32.6	1693	1	Y163_SYNY3

HEPA_HCMVA
ID HEPA_HCMVA STANDARD; PRT; 798 AA.
AC P16827;
DT 01-AUG-1990 (Rel. 15, Created)
DT 01-AUG-1990 (Rel. 15, Last sequence update)
DT 01-OCT-1996 (Rel. 34, Last annotation update)
DE DNA helicase/primase complex associated protein.
GN UL102.
OS Human cytomegalovirus (strain AD169).
OC Viruses; dsDNA viruses, no RNA stage; Herpesviridae;
OC Betaherpesvirinae; Cytomegalovirus.
OX NCBI_TaxID=10360;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=90269039; PubMed=2161319;
RA Chee M.S., Bankier A.T., Beck S., Bohni R., Brown C.M., Cerny R.,
Horsnell T., Hutchinson C.A., Ill, Kouzarides T., Martignetti J.A.,
Preddie E., Satchwell S.C., Tomlinson P., Weston K.W., Barrett B.G.,
RT "Analysis of the protein-coding content of the sequence of human
cytomegalovirus strain AD169."
RT Curr. Top. Microbiol. Immunol. 154:125-169(1990).
RL
CC -1- FUNCTION: PROBABLY INVOLVED IN DNA REPLICATION.
CC -1- SIMILARITY: BELONGS TO FAMILY THAT GROUPS TOGETHER HSV-1 UL8,
EBV-1 54, VZV 52 AND HCMV 102.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (see <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC
DR EMBL; X17403; CAA35338.1; -
DR PIR; S09867; S09867.
DR DNA replication.
KW
SQ SEQUENCE 798 AA; 85613 MW; 6A4AE14B1E1F5EB CRC64;

Query Match 97.8%; Score 132; DB 1; Length 798;
Best Local Similarity 96.2%; Pred. No. 5.7e-11;
Matches 25; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

DB 1 TWLEERDEVRSIADVAQHARRVAS 26
763 TWLEERDEVRSIADVAQHARRVAS 788

RESULT 2
ID GFAP_BOVIN STANDARD; PRT; 428 AA.
AC Q28115;
DT 01-NOV-1997 (Rel. 35, Created)
DT 01-NOV-1997 (Rel. 35, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Gial fibrillary acidic protein, astrocyte (GFAP).

```

GN GRAP.
OS Bos taurus (bovine).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eumetazoa; Cetartiodactyla; Ruminantia; Pecora; Bovidae;
OC Bovidae; Bovinae; Bos.
OX NCBI_TaxID=9913;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Subcommisural organ;
RA Bouchard P.;
RL Submitted (SEP-1996) to the EMBL/GenBank/DBJ databases.
RN [2]
RP SEQUENCE OF 197-227 FROM N.A.
RA Kikpatrick B.W., Hart G.L., Moore S.S., Barendse M.;
RL Submitted (OCT-1993) to the EMBL/GenBank/DBJ databases.
CC -1- FUNCTION: GRAP, a class-III intermediate filament, is a cell-
CC specific marker that, during the development of the central
CC nervous system, distinguishes astrocytes from other glial cells.
CC -1- SIMILARITY: Belongs to the intermediate filament family.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL; Y08255; CA659422.1; -.
CC DR EMBL; L19667; AA51413.1; -.
CC DR InterPro; IPR006821; Filament_head.
CC DR InterPro; IPR001664; IF.
CC DR InterPro; IPR002957; Keratin_1.
CC DR Pfam; PF00038; filament_1.
CC DR Pfam; PF04732; filament_head_1.
CC DR PRINTS; PR01248; TYPE1KERATIN.
CC DR PROSITE; PS00226; IF; 1.
CC KW Intermediate filament; Coiled coil.
CC FT DOMAIN 1 68 HEAD.
CC FT DOMAIN 69 373 ROD.
CC FT DOMAIN 374 428 TAIL.
CC FT DOMAIN 69 100 COIL_1A.
CC FT DOMAIN 101 111 LINKER_1.
CC FT DOMAIN 112 210 COIL_1B.
CC FT DOMAIN 211 226 LINKER_12.
CC FT DOMAIN 227 248 COIL_2A.
CC FT DOMAIN 249 252 LINKER_2.
CC FT DOMAIN 253 373 COIL_2B.
CC SQ SEQUENCE 428 AA; 49452 MW; 6891CBBF2F59AD4 CRC64;
QY Query Match 37.8%; Score 51; DB 1; Length 428;
DB Best Local Similarity 47.8%; Pred. No. 7.7;
DB Matches 11; Conservative 2; Mismatches 10; Indels 0; Gaps 0;
QY 3 LEERDENYRSLVDAQAHARRVA 25
DB 246 WHEAEWYRSKPADLNDARRNA 268

```

```

RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=94019276; PubMed=8413204;
RA Caffery R., Young P.R., McLaughlin M.W., Bergsma D.J., Koltin Y.,
RA Sathe G.M., Faucette L., Eng W.-K., Johnson R.K., Livi G.P.;
RT "Dominant missense mutations in a novel yeast protein related to
RT mammalian phosphatidylinositol 3-kinase and Vps34 abrogate rapamycin
RT cytotoxicity."
RL Mol. Cell. Biol. 13:6012-6023(1993).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=JK-3D;
RX MEDLINE=94243030; PubMed=8186460;
RA Hellmuth S.B., Wagner P., Kunz J., Deuter-Reinhard M., Henriquez R.,
RA Hall M.N.;
RT "TOR1 and TOR2 are structurally and functionally similar but not
RT identical phosphatidylinositol kinase homologues in yeast."
RL Mol. Biol. Cell 5:105-118(1994).
RN [3]
RP SEQUENCE FROM N.A.
RC STRAIN=S288C;
RX MEDLINE=96437976; PubMed=8840504;
RA Huang M.-B., Manns V., Chait J.-C., Galibert F.;
RT "Analysis of a 62 kb DNA sequence of chromosome X reveals 36 open
RT reading frames and a gene cluster with a counterpart on chromosome
RT XI."
RL Yeast 12:669-675(1996).
CC -1- FUNCTION: Phosphatidylinositol 3-kinase homolog required for G1
CC progression. Target of the antibiotic rapamycin.
CC -1- CATALYTIC ACTIVITY: ATP + 1-phosphatidy1-ID-myo-inositol = ADP +
CC 1-phosphatidy1-ID-myo-inositol 3-phosphate.
CC -1- MISCELLANEOUS: It may act on another substrate or phosphorylate a
CC different position in the phosphatidylinositol ring.
CC -1- SIMILARITY: Belongs to the p13/p14-kinase family.
CC -1- SIMILARITY: Contains 12 HEAT repeats.
CC -1- CAUTION: It is uncertain whether Met-1 is the initiator.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL; L19540; AAB66881.1; -.
CC DR EMBL; X74857; CA52849.1; -.
CC DR EMBL; Z49566; CA889594.1; -.
CC DR EMBL; L47993; AAB39292.1; -.
CC DR PIR; S57085; S57085.
CC DR HSP; P42345; IPAF.
CC DR GenOnline; 141899; -.
CC DR SGD; S0003827; TOR1.
CC DR GO; GO:0010006; C:endosome membrane; IDA.
CC DR GO; GO:0000139; C:Golgi membrane; IDA.
CC DR GO; GO:0005886; C:plasma membrane; IDA.
CC DR GO; GO:0005774; C:vacuolar membrane; IDA.
CC DR GO; GO:0005515; F:protein binding; IPI.
CC DR GO; GO:0000800; P:G1 phase of mitotic cell cycle; IMP.
CC DR GO; GO:0007126; P:meiosis; IMP.
CC DR GO; GO:0000074; P:regulation of cell cycle; IMP.
CC DR GO; GO:0007046; P:ribosome biogenesis; IMP.
CC DR InterPro; IPR008938; ARM.
CC DR InterPro; IPR003151; FAT.
CC DR InterPro; IPR003152; FATC.
CC DR InterPro; IPR000357; HEAT.
CC DR InterPro; IPR004043; P13_P14_kinase.
CC DR Pfam; PF02259; FAT; 1.
CC DR Pfam; PF02260; FATC; 1.
CC DR Pfam; PF00454; P13_P14_kinase; 1.
CC DR SMART; SM00146; PI3Kc; 1.
CC DR PROSITE; PS00077; HEAT_REPEAT; FALSE NEG.
CC DR PROSITE; PS00915; P13_P14_KINASE_1; 1.

```

```

DR PROSITE; PS00916; P13_4_KINASE_2; 1.
DR PROSITE; PS00290; P13_4_KINASE_3; 1.
KW Transferase; Kinase; Cell cycle; Repeat.
FT REPEAT 114 151
FT REPEAT 249 286
FT REPEAT 441 447
FT REPEAT 627 663
FT REPEAT 701 747
FT REPEAT 788 826
FT REPEAT 832 870
FT REPEAT 908 946
FT REPEAT 950 987
FT REPEAT 1069 1107
FT REPEAT 1109 1147
FT REPEAT 1870 1907
FT DOMAIN 2125 2470
FT MTAGEN 1972 1972
FT CONFLICT 58 58
FT CONFLICT 115 115
FT CONFLICT 133 133
FT CONFLICT 231 231
FT CONFLICT 396 396
FT CONFLICT 547 547
FT CONFLICT 675 675
FT CONFLICT 1292 1292
FT CONFLICT 1436 1436
FT CONFLICT 1468 1468
FT CONFLICT 1478 1478
FT CONFLICT 1479 1479
FT CONFLICT 1590 1590
FT CONFLICT 1632 1642
FT CONFLICT 1640 1640
FT CONFLICT 1844 1844
FT CONFLICT 1972 1972
FT CONFLICT 2202 2202
FT CONFLICT 2414 2414
SQ SEQUENCE 2470 AA; 281136 MW; ACB1781B963B81E CRC64;

Query March 37.8%; Score 51; DB 1; Length 2470;
Best Local Similarity 66.7%; Pred. No. 45;
Matches 10; Conservative 1; Mismatches 4; Indels 0; Gaps 0;

Qy 1 TWLEPDEWVSLAV 15
Db 2249 TWLEPRTYTRSLAV 2263

RESULT 4
TOR2_YEAST STANDARD; PRT; 2473 AA.
ID TOR2_YEAST
AC P33600;
DT 01-OCT-1993 (Rel. 27, Created)
DT 01-JUN-1994 (Rel. 28, Last sequence update)
DT 15-MAR-2004 (Rel. 43, Last annotation update)
DE Phosphatidylinositol 3-kinase TOR2 (EC 2.7.1.137) (P13-kinase)
DE (Pcdins-3-kinase) (PI3K).
DE TOR2 OR DRR2 OR YKL203C.
OS Saccharomyces cerevisiae (Baker's yeast).
OC Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
OC Saccharomycetales; Saccharomycetaceae; Saccharomycetes;
OC NCBI_TaxID=4932;
OX NCBI [1]
RP SEQUENCE FROM N.A.
RC STRAIN=JK3-3D;
RX MEDLINE=93258821; PubMed=8387896;
RA Kunz J., Henriquez R., Schneider U., Deuter-Reinhard M., Morva N.,
RA Hall M.N.;
RT "Target of rapamycin in yeast, TOR2, is an essential
RT phosphatidylinositol kinase homolog required for G1 progression.";
RL Cell 73:585-596(1993).
RN [2]
RP SEQUENCE OF 1-2360 FROM N.A.

```

```

RA Maia e Silva A., Bossier P., Vilela C., Fernandes L., Soares H.,
RA Guerreiro P., Rodrigues-Pousada C.;
RL Submitted (Mar-1994) to the EMBL/GenBank/DBJ databases.
RN [3]
RP SEQUENCE OF 1902-2473 FROM N.A.
RA Pohl T.M., Pohl P.M.;
RL Submitted (Mar-1994) to the EMBL/GenBank/DBJ databases.
CC -1- FUNCTION: Phosphatidylinositol 3-kinase homolog required for G1
CC progression. Target of the antibiotic rapamycin.
CC -1- CATALYTIC ACTIVITY: ATP + 1-phosphatidyL-ID-myo-inositol = ADP +
CC 1-phosphatidyL-ID-myo-inositol 3-phosphate.
CC -1- SIMILARITY: Belongs to the P13/P14-Kinase family.
CC -1- SIMILARITY: Contains 12 HEAT repeats.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (see http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL; X71416; CAA50548.1; -
CC EMBL; Z28203; CAA82048.1; -
CC PIR; S38040; S38040.
CC HSSP; P42345; 1AUB.
CC Geronline; 139957; -.
CC SGD; S0001686; TOR2.
CC GO; GO:0005886; Cytoplasm membrane; IDA.
CC GO; GO:0005774; Cytoplasmic membrane; IDA.
CC GO; GO:0005515; F-protein binding; IPI.
CC GO; GO:0007010; P-cytoskeleton organization and biogenesis; IMP.
CC GO; GO:0007266; P-Rho protein signal transduction; IMP.
CC GO; GO:0007046; P-ribosome biogenesis; IMP.
CC InterPro; IPR008938; ARM.
CC InterPro; IPR003151; FAT.
CC InterPro; IPR003152; FATC.
CC InterPro; IPR000357; HEAT.
CC InterPro; IPR000403; P13_P14_kinase.
CC InterPro; IPR008941; TPR-like.
CC Pfam; PF02259; FAT; 1.
CC Pfam; PF02260; FATC; 1.
CC Pfam; PF0454; P13_P14_kinase; 1.
CC SMART; SM00146; PI3Kc; 1.
CC DR PROSITE; PS00077; HEAT_REPEAT; FALSE_NEG.
CC DR PROSITE; PS00915; P13_4_KINASE_1; 1.
CC DR PROSITE; PS00916; P13_4_KINASE_2; 1.
CC DR PROSITE; PS00916; P13_4_KINASE_3; 1.
KW Transferase; Kinase; Cell cycle; Repeat.
FT REPEAT 588 626
FT REPEAT 636 674
FT REPEAT 676 710
FT REPEAT 710 756
FT REPEAT 756 793
FT REPEAT 797 835
FT REPEAT 841 879
FT REPEAT 917 955
FT REPEAT 1039 1076
FT REPEAT 1079 1116
FT REPEAT 1118 1155
FT REPEAT 1292 1331
FT REPEAT 1872 1909
FT REPEAT 2128 2473
FT DOMAIN 2128 2473
FT CONFLICT 1472 1472
SQ SEQUENCE 2473 AA; 281508 MW; 42C73D65CDD4DB5F CRC64;

Query March 37.8%; Score 51; DB 1; Length 2473;
Best Local Similarity 66.7%; Pred. No. 45;
Matches 10; Conservative 1; Mismatches 4; Indels 0; Gaps 0;

Qy 1 TWLEPDEWVSLAV 15
Db 2252 TWLEPRTYTRSLAV 2266

```

```

RESULT 5
OPGH_PSEAE STANDARD; PRT; 861 AA.
AC OGHU6;
DT 10-OCT-2003 (Rel. 42, Created)
DT 10-OCT-2003 (Rel. 42, Last sequence update)
DE 10-OCT-2003 (Rel. 42, Last annotation update)
DE Glucans biosynthesis glucosyltransferase H (EC 2.4.1.-).
GN OPGH OR MOOH OR PA5077.
OS Pseudomonas aeruginosa.
OC Bacteria; Proteobacteria; Gammaproteobacteria; Pseudomonadales;
OC Pseudomonadaceae; Pseudomonas.
OK NCBI_TaxID=287;
RN
RP SEQUENCE FROM N.A.
RC STRAIN=ATCC 15692 / PA01;
RX MEDLINE=20437337; PubMed=10984043;
RA Hickey M.J., Brinkman F.S.L., Hufnagle W.O., Kowalik D.J., Laerou M.,
RA Gardner R.L., Goltzer L., Tolentino E., Westbrook-Wadman S., Yuan Y.,
RA Brody L.L., Coulter S.N., Folger K.R., Kas A., Lardig K., Lim R.M.,
RA Smith K.A., Spencer D.H., Wong G.K.-S., Wu Z., Paulsen I.T.,
RA Reizer J., Sailer M.H., Hancock R.E.W., Lory S., Olson M.V.;
RT "Complete genome sequence of Pseudomonas aeruginosa PA01, an
RT opportunistic pathogen.";
RL Nature 406:959-964(2000).
CC -1- FUNCTION: Involved in the biosynthesis of osmoregulated
CC periplasmic glucans (OPGs) (by similarity).
CC -1- PATHWAY: OPG (osmoregulated periplasmic glucans) biosynthesis.
CC -1- SUBCELLULAR LOCATION: Integral membrane protein. Inner membrane
CC (by similarity).
CC -1- SIMILARITY: Belongs to the glucosyltransferase family 2. OpgH
CC subfamily.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC
CC -----
CC DR EMBL; AE004920; AACG08462.1; -.
CC DR PIR; H83012; H83012.
CC DR HAMAP; MF 01072; -.
CC DR InterPro; IPR001173; GLYCO.trans.2.
CC DR Pfam; PF00535; Glycosyltransf.2; I.
CC KW Transferase; Glycosyltransferase; Inner membrane; Transmembrane;
CC Complete proteome.
CC FT TRANSMEM 142 159 Potential.
CC FT TRANSMEM 198 220 Potential.
CC FT TRANSMEM 521 543 Potential.
CC FT TRANSMEM 572 594 Potential.
CC FT TRANSMEM 601 623 Potential.
CC FT TRANSMEM 683 705 Potential.
CC SQ SEQUENCE 861 AA; 96994 MW; 63557B5A63107E7E CRC64;
Query Match 37.4%; Score 50.5; DB 1; Length 861;
Best Local Similarity 43.5%; Pred. No. 18;
Matches 10; Conservative 3; Mismatches 7; Indels 3; Gaps 1;

```

```

DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Preprotein translocase secA 1 subunit.
GN SECAL OR SECA OR ML0779.
OS Mycobacterium lepreae.
OC Bacteria; Actinobacteria; Actinobacteriales;
OC Corynebacteriaceae; Mycobacteriaceae; Mycobacterium.
OK NCBI_TaxID=1769;
RN
RP SEQUENCE FROM N.A.
RC STRAIN=TN;
RX MEDLINE=21128732; PubMed=11234002;
RA Cole S.T., Eiglmister K., Parkhill J., James K.D., Thomson N.R.,
RA Wheeler P.R., Honore N., Garnier T., Churruarin C., Harrie D.,
RA Mungall K., Basham D., Brown D., Chillingworth T., Connor R.,
RA Davies R.M., Devlin K., Duthoy S., Feltham T., Fraser A., Hamlin N.,
RA Holroyd S., Hornsby T., Jagels K., Lacroix C., Maclean J., Moule S.,
RA Murthy L., Oliver K., Quail M.A., Rajandream M.A., Rutherford K.M.,
RA Rutter S., Seeger K., Simon S., Simmonds M., Skelton J., Squares R.,
RA Squares S., Stevens K., Taylor K., Whitehead S., Woodward J.R.,
RA Barrett B.G.;
RT "Massive gene decay in the leprosy bacillus.";
RL Nature 409:1007-1011(2001).
CC -1- FUNCTION: Involved in protein export. Interacts with the secY/secE
CC subunits. SecA has a central role in coupling the hydrolysis of
CC ATP to the transfer of pre-secretory periplasmic and outer
CC membrane proteins across the membrane (by similarity).
CC -1- SUBUNIT: Part of the prokaryotic protein translocation apparatus
CC which comprise secA, secB, secE, secY, secZ and secE (by
CC similarity).
CC -1- SUBCELLULAR LOCATION: Cytoplasmic side of membrane (by
CC similarity).
CC -1- SIMILARITY: Belongs to the secA family.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC
CC -----
CC DR EMBL; AL583919; CAC30288.1; -.
CC DR PIR; D87006; D87006.
CC DR Leproma; ML0779; -.
CC DR InterPro; IPR000185; SecA.
CC DR Pfam; PF01043; SecA protein; 1.
CC DR PRINTS; PR00906; SECA.
CC DR TIGRfam; TIGR00963; secA; 1.
CC DR PROSITE; PS01812; SECA; 1.
CC KW Protein transport; ATP-binding; Membrane; Translocation; Transport;
CC Complete proteome.
CC FT NP BIND 101 108 ATP (POTENTIAL).
CC SQ SEQUENCE 940 AA; 105528 MW; 00AABBF62D76DD71 CRC64;
Query Match 36.3%; Score 49; DB 1; Length 940;
Best Local Similarity 60.0%; Pred. No. 33;
Matches 9; Conservative 3; Mismatches 3; Indels 0; Gaps 0;

```

```

RESULT 6
SEAL_MYCLE STANDARD; PRT; 940 AA.
AC PS7936;
DT 16-OCT-2001 (Rel. 40, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)

```

```

RESULT 7
ID SPCC RAT STANDARD; PRT; 2388 AA.
AC O90WN8; O88197; O98568;
DT 28-FEB-2003 (Rel. 41, Created)
DT 28-FEB-2003 (Rel. 41, Last sequence update)
DT 10-OCT-2003 (Rel. 42, Last annotation update)
DE Spectrin beta chain, brain 2 (Spectrin, non-erythroid beta chain 2)
DE (Beta-III spectrin) (SPNB-3) (Beta Spili sigma 1) (Spectrin-like
DE protein GRPAP41).

```

CN SPTBN2.
 OS Rattus norvegicus (Rat).
 CC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 CC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
 CX NCBI_TaxID=10116;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=Sprague-Dawley; TISSUE=Brain;
 RX MEDLINE=98340016; PubMed=9675416;
 RA Ohara O., Ohara R., Yamakawa H., Nakajima D., Nakayama M.;
 RT "Characterization of a new beta-spectrin gene which is predominantly
 expressed in brain";
 RL Brain Res. Mol. Brain Res. 57:181-192(1998).
 RN [2]
 RP SEQUENCE FROM N.A.
 RC STRAIN=Sprague-Dawley; TISSUE=Brain;
 RX MEDLINE=98369635; PubMed=9704016;
 RA Sakaguchi G., Orita S., Naito A., Maeda M., Igarashi H., Sasaki T.,
 RA Takai Y.;
 RT "A novel brain-specific isoform of beta spectrin: isolation and its
 interaction with Munc13";
 RL Biochem. Biophys. Res. Commun. 248:846-851(1998).
 RN [3]
 RP SEQUENCE FROM N.A.
 RC Rochstein J.D., Jackson M.;
 RX "GTRAP41 - a cytoskeletal modulator of EAAT4 glutamate transporter";
 RT Submitted (JAN-2000) to the EMBL/GenBank/DBJ databases.
 CC -1- FUNCTION: Probably plays an important role in neuronal membrane
 skeleton.
 CC -1- TISSUE SPECIFICITY: Abundantly transcribed in the brain. Neurons
 are the predominant cell-type to express the gene. Found
 abundantly in Purkinje cells.
 CC -1- SIMILARITY: Belongs to the spectrin family.
 CC -1- SIMILARITY: Contains 2 calponin-homology (CH) domains.
 CC -1- SIMILARITY: Contains 1 PH domain.
 CC -1- SIMILARITY: Contains 17 spectrin repeats.
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 between the Swiss Institute of Bioinformatics and the EMBL outstation -
 the European Bioinformatics Institute. There are no restrictions on its
 use by non-profit institutions as long as its content is in no way
 modified and this statement is not removed. Usage by and for commercial
 entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 or send an email to license@isb-sib.ch).
 CC -----
 CC EMBL; AB008551; BAA32699.1; -
 CC EMBL; AB001347; BAA32473.1; -
 CC EMBL; AF225960; AAC28596.1; -
 CC PIR; J0271; J0271.
 CC HSSP; Q01082; IBKR.
 DR GO; GO:0005737; C:cytoplasm; TAS.
 DR GO; GO:0016363; C:nuclear matrix; TAS.
 DR GO; GO:0008091; C:spectrin; TAS.
 DR GO; GO:0003779; F:actin binding; ISS.
 DR GO; GO:0030506; F:ankyrin binding; ISS.
 DR GO; GO:0005200; F:structural constituent of cytoskeleton; TAS.
 DR GO; GO:0017158; P:regulation of calcium ion dependent exocytosis; TAS.
 DR GO; GO:0016081; P:synaptic vesicle docking; TAS.
 DR InterPro; IPR001589; Actbind-actin.
 DR InterPro; IPR001715; Calponin-like.
 DR InterPro; IPR001849; PH.
 DR InterPro; IPR002017; Spectrin.
 DR InterPro; IPR001605; Spectrin_PH.
 DR Pfam; PF00307; CH; 2.
 DR Pfam; PF00169; PH; 1.
 DR Pfam; PF00435; Spectrin; 17.
 DR PRINTS; PRO0683; SPECTRINPH.
 DR SMART; SM00033; CH; 2.
 DR SMART; SM00233; PH; 1.
 DR SMART; SM00150; SPEC; 16.
 DR PROSITE; PS00019; ACTININ_1; 1.
 DR PROSITE; PS00020; ACTININ_2; 1.
 DR PROSITE; PS50021; CH; 2.

DR PROSITE; PS50003; PH DOMAIN; 1.
 KM Cytoskeleton; Membrane; Repeat; Actin-binding; Actin capping.
 FT DOMAIN 1 278
 FT DOMAIN 1 278
 FT DOMAIN 1 278
 FT DOMAIN 1 278
 FT REPEAT 176 278
 FT REPEAT 305 415
 FT REPEAT 425 529
 FT REPEAT 531 639
 FT REPEAT 641 745
 FT REPEAT 747 850
 FT REPEAT 852 956
 FT REPEAT 958 1063
 FT REPEAT 1065 1170
 FT REPEAT 1172 1276
 FT REPEAT 1278 1381
 FT REPEAT 1383 1486
 FT REPEAT 1488 1588
 FT REPEAT 1588 1692
 FT REPEAT 1694 1799
 FT REPEAT 1801 1905
 FT REPEAT 1907 2011
 FT REPEAT 2013 2075
 FT DOMAIN 2218 2328
 FT CONFLICT 326 543
 FT CONFLICT 543 543
 FT CONFLICT 608 608
 FT CONFLICT 887 887
 FT CONFLICT 908 908
 FT CONFLICT 948 948
 FT CONFLICT 1156 1157
 FT CONFLICT 1194 1194
 FT CONFLICT 1555 1555
 FT CONFLICT 1769 1769
 SQ SEQUENCE 2388 AA; 271062 MW; 3EC8965AF0655F19 CRC64;
 Query Match 36.3%; Score 49; DB 1; Length 2388;
 Best Local Similarity 42.9%; Pred. No. 84;
 Matches 6; Conservative 6; Mismatches 2; Indels 0; Gaps 0;
 Qy 2 WHEREDEWYSLAV 15
 Db 867 WVEKEQWNGIAL 880
 RESULT 8
 SPCF HUMAN STANDARD; PRT; 2390 AA.
 ID SPCF HUMAN 015020; 014872; 014873;
 AC 015020; 014872; 014873;
 DT 28-FEB-2003 (Rel. 41; Created)
 DT 28-FEB-2003 (Rel. 41; Last sequence update)
 DT 10-OCT-2003 (Rel. 42; Last annotation update)
 DE Spectrin beta chain, brain 2 (Spectrin, non-erythroid beta chain 2)
 DE (beta-III spectrin).
 GN SPTBN2 OR KIA0302.
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
 CX NCBI_TaxID=9606;
 RN [1]
 RP SEQUENCE FROM N.A. (ISOFORM 1).
 RC TISSUE=Brain;
 RX MEDLINE=97349984; PubMed=9205841;
 RA Nagase T., Ishikawa K.-I., Nakajima D., Ohira M., Seki N.,
 RA Miyajima N., Tanaka A., Kotani H., Nomura N., Ohara O.;
 RT "Prediction of the coding sequences of unidentified human genes. VII.
 RT The complete sequences of 100 new cDNA clones from brain which can
 RT code for large proteins in vitro";
 RL DNA Res. 4:141-150(1997).
 RN [2]
 RP SEQUENCE OF 1-34 FROM N.A.
 RC TISSUE=Brain;
 RX MEDLINE=99045654; PubMed=9826670;
 RA Stanekwich M.C., Tse W.T., Peters L.L., Ch'ng Y., John K.M.,

FT	REPEAT	852	956	SPECTRIN 6.
FT <th>REPEAT</th> <td>958<td>1063<th>SPECTRIN 7.</th></td></td>	REPEAT	958 <td>1063<th>SPECTRIN 7.</th></td>	1063 <th>SPECTRIN 7.</th>	SPECTRIN 7.
FT <th>REPEAT</th> <td>1065<td>1170<th>SPECTRIN 8.</th></td></td>	REPEAT	1065 <td>1170<th>SPECTRIN 8.</th></td>	1170 <th>SPECTRIN 8.</th>	SPECTRIN 8.
FT <th>REPEAT</th> <td>1172<td>1276<th>SPECTRIN 9.</th></td></td>	REPEAT	1172 <td>1276<th>SPECTRIN 9.</th></td>	1276 <th>SPECTRIN 9.</th>	SPECTRIN 9.
FT <th>REPEAT</th> <td>1278<td>1381<th>SPECTRIN 10.</th></td></td>	REPEAT	1278 <td>1381<th>SPECTRIN 10.</th></td>	1381 <th>SPECTRIN 10.</th>	SPECTRIN 10.
FT <th>REPEAT</th> <td>1383<td>1486<th>SPECTRIN 11.</th></td></td>	REPEAT	1383 <td>1486<th>SPECTRIN 11.</th></td>	1486 <th>SPECTRIN 11.</th>	SPECTRIN 11.
FT <th>REPEAT</th> <td>1488<td>1586<th>SPECTRIN 12.</th></td></td>	REPEAT	1488 <td>1586<th>SPECTRIN 12.</th></td>	1586 <th>SPECTRIN 12.</th>	SPECTRIN 12.
FT <th>REPEAT</th> <td>1588<td>1692<th>SPECTRIN 13.</th></td></td>	REPEAT	1588 <td>1692<th>SPECTRIN 13.</th></td>	1692 <th>SPECTRIN 13.</th>	SPECTRIN 13.
FT <th>REPEAT</th> <td>1694<td>1799<th>SPECTRIN 14.</th></td></td>	REPEAT	1694 <td>1799<th>SPECTRIN 14.</th></td>	1799 <th>SPECTRIN 14.</th>	SPECTRIN 14.
FT <th>REPEAT</th> <td>1801<td>1905<th>SPECTRIN 15.</th></td></td>	REPEAT	1801 <td>1905<th>SPECTRIN 15.</th></td>	1905 <th>SPECTRIN 15.</th>	SPECTRIN 15.
FT <th>REPEAT</th> <td>1907<td>2011<th>SPECTRIN 16.</th></td></td>	REPEAT	1907 <td>2011<th>SPECTRIN 16.</th></td>	2011 <th>SPECTRIN 16.</th>	SPECTRIN 16.
FT <th>REPEAT</th> <td>2013<td>2075<th>SPECTRIN 17.</th></td></td>	REPEAT	2013 <td>2075<th>SPECTRIN 17.</th></td>	2075 <th>SPECTRIN 17.</th>	SPECTRIN 17.
FT <th>DOMAIN</th> <td>2218<td>2328<th>PH.</th></td></td>	DOMAIN	2218 <td>2328<th>PH.</th></td>	2328 <th>PH.</th>	PH.
FT <th>VARSPLIC</th> <td>2214<td>2390<th>PH.</th></td></td>	VARSPLIC	2214 <td>2390<th>PH.</th></td>	2390 <th>PH.</th>	PH.
FT <td></td> <td></td> <td></td> <td>AEHMSYLRVNAATAATSSASGEPEEPVSTTRGMRMT</td>				AEHMSYLRVNAATAATSSASGEPEEPVSTTRGMRMT
FT <td></td> <td></td> <td></td> <td>MPSPSGVAGPGVTLREKDRERERERERFFKXNK -> V</td>				MPSPSGVAGPGVTLREKDRERERERERFFKXNK -> V
FT <td></td> <td></td> <td></td> <td>SCBSCGSLVSPGQPLPADSPSPVPLPFFGLVLCGRTGCV</td>				SCBSCGSLVSPGQPLPADSPSPVPLPFFGLVLCGRTGCV
FT <td></td> <td></td> <td></td> <td>RRHQALPV (In Isoform 2).</td>				RRHQALPV (In Isoform 2).
FT <td></td> <td></td> <td></td> <td>/FtId=VSP_000722.</td>				/FtId=VSP_000722.
FT <td></td> <td></td> <td></td> <td>SPECTRIN 16.</td>				SPECTRIN 16.
FT <td></td> <td></td> <td></td> <td>SPECTRIN 17.</td>				SPECTRIN 17.
FT <td></td> <td></td> <td></td> <td>SPECTRIN 18.</td>				SPECTRIN 18.
FT <td></td> <td></td> <td></td> <td>SPECTRIN 19.</td>				SPECTRIN 19.
FT <td></td> <td></td> <td></td> <td>SPECTRIN 20.</td>				SPECTRIN 20.
FT <td></td> <td></td> <td></td> <td>SPECTRIN 21.</td>				SPECTRIN 21.
FT <td></td> <td></td> <td></td> <td>SPECTRIN 22.</td>				SPECTRIN 22.
FT <td></td> <td></td> <td></td> <td>SPECTRIN 23.</td>				SPECTRIN 23.
FT <td></td> <td></td> <td></td> <td>SPECTRIN 24.</td>				SPECTRIN 24.
FT <td></td> <td></td> <td></td> <td>SPECTRIN 25.</td>				SPECTRIN 25.
FT <td></td> <td></td> <td></td> <td>SPECTRIN 26.</td>				SPECTRIN 26.
FT <td></td> <td></td> <td></td> <td>SPECTRIN 27.</td>				SPECTRIN 27.
FT <td></td> <td></td> <td></td> <td>SPECTRIN 28.</td>				SPECTRIN 28.
FT <td></td> <td></td> <td></td> <td>SPECTRIN 29.</td>				SPECTRIN 29.
FT <td></td> <td></td> <td></td> <td>SPECTRIN 30.</td>				SPECTRIN 30.
FT <td></td> <td></td> <td></td> <td>SPECTRIN 31.</td>				SPECTRIN 31.
FT <td></td> <td></td> <td></td> <td>SPECTRIN 32.</td>				SPECTRIN 32.
FT <td></td> <td></td> <td></td> <td>SPECTRIN 33.</td>				SPECTRIN 33.
FT <td></td> <td></td> <td></td> <td>SPECTRIN 34.</td>				SPECTRIN 34.
FT <td></td> <td></td> <td></td> <td>SPECTRIN 35.</td>				SPECTRIN 35.
FT <td></td> <td></td> <td></td> <td>SPECTRIN 36.</td>				SPECTRIN 36.
FT <td></td> <td></td> <td></td> <td>SPECTRIN 37.</td>				SPECTRIN 37.
FT <td></td> <td></td> <td></td> <td>SPECTRIN 38.</td>				SPECTRIN 38.
FT <td></td> <td></td> <td></td> <td>SPECTRIN 39.</td>				SPECTRIN 39.
FT <td></td> <td></td> <td></td> <td>SPECTRIN 40.</td>				SPECTRIN 40.
FT <td></td> <td></td> <td></td> <td>SPECTRIN 41.</td>				SPECTRIN 41.
FT <td></td> <td></td> <td></td> <td>SPECTRIN 42.</td>				SPECTRIN 42.
FT <td></td> <td></td> <td></td> <td>SPECTRIN 43.</td>				SPECTRIN 43.
FT <td></td> <td></td> <td></td> <td>SPECTRIN 44.</td>				SPECTRIN 44.
FT <td></td> <td></td> <td></td> <td>SPECTRIN 45.</td>				SPECTRIN 45.
FT <td></td> <td></td> <td></td> <td>SPECTRIN 46.</td>				SPECTRIN 46.
FT <td></td> <td></td> <td></td> <td>SPECTRIN 47.</td>				SPECTRIN 47.
FT <td></td> <td></td> <td></td> <td>SPECTRIN 48.</td>				SPECTRIN 48.
FT <td></td> <td></td> <td></td> <td>SPECTRIN 49.</td>				SPECTRIN 49.
FT <td></td> <td></td> <td></td> <td>SPECTRIN 50.</td>				SPECTRIN 50.
FT <td></td> <td></td> <td></td> <td>SPECTRIN 51.</td>				SPECTRIN 51.
FT <td></td> <td></td> <td></td> <td>SPECTRIN 52.</td>				SPECTRIN 52.
FT <td></td> <td></td> <td></td> <td>SPECTRIN 53.</td>				SPECTRIN 53.
FT <td></td> <td></td> <td></td> <td>SPECTRIN 54.</td>				SPECTRIN 54.
FT <td></td> <td></td> <td></td> <td>SPECTRIN 55.</td>				SPECTRIN 55.
FT <td></td> <td></td> <td></td> <td>SPECTRIN 56.</td>				SPECTRIN 56.
FT <td></td> <td></td> <td></td> <td>SPECTRIN 57.</td>				SPECTRIN 57.
FT <td></td> <td></td> <td></td> <td>SPECTRIN 58.</td>				SPECTRIN 58.
FT <td></td> <td></td> <td></td> <td>SPECTRIN 59.</td>				SPECTRIN 59.
FT <td></td> <td></td> <td></td> <td>SPECTRIN 60.</td>				SPECTRIN 60.
FT <td></td> <td></td> <td></td> <td>SPECTRIN 61.</td>				SPECTRIN 61.
FT <td></td> <td></td> <td></td> <td>SPECTRIN 62.</td>				SPECTRIN 62.
FT <td></td> <td></td> <td></td> <td>SPECTRIN 63.</td>				SPECTRIN 63.
FT <td></td> <td></td> <td></td> <td>SPECTRIN 64.</td>				SPECTRIN 64.
FT <td></td> <td></td> <td></td> <td>SPECTRIN 65.</td>				SPECTRIN 65.
FT <td></td> <td></td> <td></td> <td>SPECTRIN 66.</td>				SPECTRIN 66.
FT <td></td> <td></td> <td></td> <td>SPECTRIN 67.</td>				SPECTRIN 67.
FT <td></td> <td></td> <td></td> <td>SPECTRIN 68.</td>				SPECTRIN 68.
FT <td></td> <td></td> <td></td> <td>SPECTRIN 69.</td>				SPECTRIN 69.
FT <td></td> <td></td> <td></td> <td>SPECTRIN 70.</td>				SPECTRIN 70.
FT <td></td> <td></td> <td></td> <td>SPECTRIN 71.</td>				SPECTRIN 71.
FT <td></td> <td></td> <td></td> <td>SPECTRIN 72.</td>				SPECTRIN 72.
FT <td></td> <td></td> <td></td> <td>SPECTRIN 73.</td>				SPECTRIN 73.
FT <td></td> <td></td> <td></td> <td>SPECTRIN 74.</td>				SPECTRIN 74.
FT <td></td> <td></td> <td></td> <td>SPECTRIN 75.</td>				SPECTRIN 75.
FT <td></td> <td></td> <td></td> <td>SPECTRIN 76.</td>				SPECTRIN 76.
FT <td></td> <td></td> <td></td> <td>SPECTRIN 77.</td>				SPECTRIN 77.
FT <td></td> <td></td> <td></td> <td>SPECTRIN 78.</td>				SPECTRIN 78.
FT <td></td> <td></td> <td></td> <td>SPECTRIN 79.</td>				SPECTRIN 79.
FT <td></td> <td></td> <td></td> <td>SPECTRIN 80.</td>				SPECTRIN 80.
FT <td></td> <td></td> <td></td> <td>SPECTRIN 81.</td>				SPECTRIN 81.
FT <td></td> <td></td> <td></td> <td>SPECTRIN 82.</td>				SPECTRIN 82.
FT <td></td> <td></td> <td></td> <td>SPECTRIN 83.</td>				SPECTRIN 83.
FT <td></td> <td></td> <td></td> <td>SPECTRIN 84.</td>				SPECTRIN 84.
FT <td></td> <td></td> <td></td> <td>SPECTRIN 85.</td>				SPECTRIN 85.
FT <td></td> <td></td> <td></td> <td>SPECTRIN 86.</td>				SPECTRIN 86.
FT <td></td> <td></td> <td></td> <td>SPECTRIN 87.</td>				SPECTRIN 87.
FT <td></td> <td></td> <td></td> <td>SPECTRIN 88.</td>				SPECTRIN 88.
FT <td></td> <td></td> <td></td> <td>SPECTRIN 89.</td>				SPECTRIN 89.
FT <td></td> <td></td> <td></td> <td>SPECTRIN 90.</td>				SPECTRIN 90.
FT <td></td> <td></td> <td></td> <td>SPECTRIN 91.</td>				SPECTRIN 91.
FT <td></td> <td></td> <td></td> <td>SPECTRIN 92.</td>				SPECTRIN 92.
FT <td></td> <td></td> <td></td> <td>SPECTRIN 93.</td>				SPECTRIN 93.
FT <td></td> <td></td> <td></td> <td>SPECTRIN 94.</td>				SPECTRIN 94.
FT <td></td> <td></td> <td></td> <td>SPECTRIN 95.</td>				SPECTRIN 95.
FT <td></td> <td></td> <td></td> <td>SPECTRIN 96.</td>				SPECTRIN 96.
FT <td></td> <td></td> <td></td> <td>SPECTRIN 97.</td>				SPECTRIN 97.
FT <td></td> <td></td> <td></td> <td>SPECTRIN 98.</td>				SPECTRIN 98.
FT <td></td> <td></td> <td></td> <td>SPECTRIN 99.</td>				SPECTRIN 99.
FT <td></td> <td></td> <td></td> <td>SPECTRIN 100.</td>				SPECTRIN 100.
FT <td></td> <td></td> <td></td> <td>SPECTRIN 101.</td>				SPECTRIN 101.
FT <td></td> <td></td> <td></td> <td>SPECTRIN 102.</td>				SPECTRIN 102.
FT <td></td> <td></td> <td></td> <td>SPECTRIN 103.</td>				SPECTRIN 103.
FT <td></td> <td></td> <td></td> <td>SPECTRIN 104.</td>				SPECTRIN 104.
FT <td></td> <td></td> <td></td> <td>SPECTRIN 105.</td>				SPECTRIN 105.
FT <td></td> <td></td> <td></td> <td>SPECTRIN 106.</td>				SPECTRIN 106.

RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
 RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
 RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,
 RA Butterfield V.S.N., Krzywinski M.I., Skalska U., Smalhus D.E.,
 RA Schnerch A., Schein J.E., Jones S.J.M., Maitra M.A.,
 RT "Generation and initial analysis of more than 15,000 full-length
 human and mouse cDNA sequences.",
 RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
 RN [4]
 RP STRUCTURE BY NMR OF 1-105.
 RX MEDLINE=94268557; PubMed=8208296;
 RA Yoon H.S., Hajduk P.J., Petros A.M., Olejniczak E.T., Meadows R.P.,
 RA Pesik S.W.,
 RT "Solution structure of a pleckstrin-homology domain,"
 RL Nature 369:672-675(1994).
 CC -!- FUNCTION: Major protein kinase C substrate of platelets, its
 CC exact function is not known.
 CC -!- SIMILARITY: Contains 1 DEP domain.
 CC -!- SIMILARITY: Contains 2 PH domains.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL; X07743; CA30564.1; -.
 DR EMBL; BC018549; AAH18549.1; -.
 DR PIR; S00755; S00755.
 DR PDB; 1PLS; 03-JUN-95.
 DR Genew; HGNC:3070; PLEK.
 DR MIM; 173570; -.
 DR GO; GO:0005509; F:calcium ion binding; TAS.
 DR InterPro; IPR000591; DEP.
 DR InterPro; IPR001849; PH.
 DR Pfam; PF00610; DEP; 1.
 DR SMART; SM00169; PH; 2.
 DR SMART; SM00233; PH; 2.
 DR PROSITE; PS50186; DEP; 1.
 DR PROSITE; PS50003; PH DOMAIN; 2.
 KW Phosphorylation; Repeat; 3D-structure.
 FT DOMAIN 4 101
 FT DOMAIN 136 221
 FT DOMAIN 244 347
 FT VARIANT 92 92
 FT -----
 FT CONFLICT 97 97
 FT STRAND 7 14
 FT STRAND 22 29
 FT TURN 30 31
 FT STRAND 32 36
 FT TURN 39 40
 FT STRAND 46 49
 FT STRAND 55 56
 FT STRAND 68 73
 FT TURN 74 76
 FT STRAND 77 82
 FT TURN 86 103
 FT TURN 104 105
 SQ SEQUENCE 350 AA; 40082 MW; 2E2A128C8B526361 CRC64;
 Query Match 34.8%; Score 47; DB 1; Length 350;
 Best Local Similarity 66.7%; Pred. No. 23;
 Matches 8; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

RESULT 12
 PLEK MOUSE
 ID PLEK MOUSE STANDARD; PRT; 350 AA.
 AC Q9JH5; O9ER19;
 DT 28-FEB-2003 (Rel. 41, Created)
 DT 28-FEB-2003 (Rel. 41, Last sequence update)
 DT 28-FEB-2003 (Rel. 41, Last annotation update)
 DE Pleckstrin.
 GN PLEK.
 OS Mus musculus (Mouse).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 OX NCBI_Taxid=10090;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=BA1B/C;
 RX MEDLINE=20318622; PubMed=10860665;
 RA Cedrik J.L., Hegamyer G., Gerrard B., Dean M., Colburn N.H.;
 RT "cDNA cloning and mapping of mouse pleckstrin (Plek), a gene
 RT upregulated in transformation-resistant cells,"
 RL Genomics 66:204-212(2000).
 RL [2]
 RP SEQUENCE FROM N.A.
 RA Zhang Y., Wu G., Palge C.J.;
 RT "Involvement of pleckstrin in B cell differentiation and
 RT activation,"
 RL Submitted (JUN-1998) to the EMBL/GenBank/DBJ databases.
 RN [3]
 RP SEQUENCE FROM N.A.
 RA Ahn H.-J., Cho J.-J.;
 RT "Mouse pleckstrin 1 is induced in mast cells after Ige cross-
 RT linking";
 RL Submitted (SEP-2000) to the EMBL/GenBank/DBJ databases.
 CC -!- FUNCTION: Major protein kinase C substrate of platelets, its exact
 CC function is not known.
 CC -!- SIMILARITY: Contains 1 DEP domain.
 CC -!- SIMILARITY: Contains 2 PH domains.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL; AF181829; AAF75830.1; -.
 DR EMBL; AF073294; AAF72039.1; -.
 DR EMBL; AF303745; AAG29513.1; -.
 DR HSSP; P08567; 1PLS.
 DR MGD; MGI:1860485; Plek.
 DR InterPro; IPR000591; DEP.
 DR InterPro; IPR001849; PH.
 DR Pfam; PF00610; DEP; 1.
 DR SMART; SM00169; PH; 2.
 DR SMART; SM00049; DEP; 1.
 DR SMART; SM00233; PH; 2.
 DR PROSITE; PS50186; DEP; 1.
 DR PROSITE; PS50003; PH DOMAIN; 2.
 KW Phosphorylation; Repeat.
 FT DOMAIN 4 101
 FT DOMAIN 136 221
 FT DOMAIN 244 347
 FT CONFLICT 120 120
 FT CONFLICT 225 225
 SQ SEQUENCE 350 AA; 39900 MW; 348F3CB469B7CC53 CRC64;
 Query Match 34.8%; Score 47; DB 1; Length 350;
 Best Local Similarity 66.7%; Pred. No. 23;
 Matches 8; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

QY 3 LEERDEWYRSIAVDAQHAAR 23
 DB 277 LOEAEBWYKSKRADLSEANR 297

RESULT 15
 VIME_MESAU STANDARD; PRT; 464 AA.
 AC P02544;
 DT 21-JUL-1986 (Rel. 01, Created)
 DT 21-JUL-1986 (Rel. 01, Last sequence update)
 DT 28-FEB-2003 (Rel. 41, Last annotation update)
 DE Vimentin.
 GN VIM.
 OS Mesocricetus auratus (Golden hamster).
 CC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 CC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Cricetinae;
 CC Mesocricetus.
 OC NCBI_TaxID=10036;
 OX 1;
 RN RP
 RP SEQUENCE FROM N.A.
 RC TISSUE=Lens;
 RX MEDLINE=84026520; PubMed=6194898;
 RA Quax W.J., Egberts W.V., Hendriks W., Quax-Deuker Y.E.F.M.,
 RT Bloemendaal H.;
 RL "The structure of the vimentin gene.";
 RL Cell 35:215-223(1983).
 [2]
 RP SEQUENCE FROM N.A.
 RC TISSUE=Lens;
 RX MEDLINE=83221633; PubMed=6304716;
 RA Quax-Deuker Y.E.F.M., Quax W.J., Bloemendaal H.;
 RT "Primary and secondary structure of hamster vimentin predicted from
 the nucleotide sequence.";
 RL Proc. Natl. Acad. Sci. U.S.A. 80:3548-3552(1983).
 CC -1- FUNCTION: Vimentins are class-III intermediate filaments found in
 various non-epithelial cells, especially mesenchymal cells.
 CC -1- SUBUNIT: Homopolymer.
 CC -1- PTM: One of the most prominent phosphoproteins in various cells of
 mesenchymal origin. Phosphorylation is enhanced during cell
 division, at which time vimentin filaments are significantly
 reorganized.
 CC -1- SIMILARITY: Belongs to the intermediate filament family.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 between the Swiss Institute of Bioinformatics and the EMBL outstation -
 the European Bioinformatics Institute. There are no restrictions on its
 use by non-profit institutions as long as its content is in no way
 modified and this statement is not removed. Usage by and for commercial
 entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL; K00927; AAA37104.1; -;
 DR EMBL; K00921; AAA37104.1; JOINED.
 DR EMBL; K00922; AAA37104.1; JOINED.
 DR EMBL; K00923; AAA37104.1; JOINED.
 DR EMBL; K00924; AAA37104.1; JOINED.
 DR EMBL; K00925; AAA37104.1; JOINED.
 DR EMBL; K00926; AAA37104.1; JOINED.
 DR PIR; A90842; VEH;
 DR InterPro; IPR006821; Filament_head.
 DR InterPro; IPR001664; IF.
 DR InterPro; IPR02957; Keratin_I.
 DR Pfam; PF00038; filament_1.
 DR Pfam; PF04732; filament_head_1.
 DR PRINTS; PR01248; TYPE1KERATIN.
 DR PROSITE; PS00226; IF; 1.
 KW Intermediate filament; Coiled coil; Acetylation; Phosphorylation.
 FT INIT MET 0
 FT DOMAIN 1 93 HEAD.
 FT DOMAIN 94 405 ROD.
 FT DOMAIN 406 464 TAIL.

FT DOMAIN 94 129 COIL 1A.
 FT DOMAIN 130 151 LINKER 1.
 FT DOMAIN 152 243 COIL 1B.
 FT DOMAIN 244 266 LINKER 12.
 FT DOMAIN 267 405 COIL 2.
 FT MOD_RES 1 1 ACETYLATION.
 FT MOD_RES 38 38 PHOSPHORYLATION (BY CAMK2) (BY
 FT MOD_RES 81 81 SIMILARITY).
 FT MOD_RES 81 81 PHOSPHORYLATION (BY CAMK2) (BY
 FT MOD_RES 81 81 SIMILARITY).
 FT CONFLICT 42 42 L -> A (IN REF. 2).
 FT CONFLICT 115 115 Y -> D (IN REF. 2).
 FT CONFLICT 182 182 R -> I (IN REF. 2).
 SQ SEQUENCE 464 AA; 53598 MW; 7AC417083CBD4776 CRC64;

Query Match 34.8%; Score 47; DB 1; Length 464;
 Best Local Similarity 42.9%; Pred. No. 30;
 Matches 9; Conservative 3; Mismatches 9; Indels 0; Gaps 0;

QY 3 LEERDEWYRSIAVDAQHAAR 23
 DB 282 LOEAEBWYKSKRADLSEANR 302

Search completed: September 30, 2004, 11:10:35
 Job time : 11.6571 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: September 30, 2004, 10:58:43 ; Search time 67.3524 Seconds
(without alignments)
109.072 Million cell updates/sec

Title: US-09-988-851A-12

Perfect score: 135
Sequence: 1 TWLEERDEWVRSIAVDAQHAAKRVAS 26

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 1586107 seqs, 282547505 residues

Total number of hits satisfying chosen parameters: 1586107

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%

Listing first 45 summaries

Database : A Geneseq_29Jan04:*
1: Geneseqp1980s:*
2: Geneseqp1990s:*
3: Geneseqp2000s:*
4: Geneseqp2001s:*
5: Geneseqp2002s:*
6: Geneseqp2003as:*
7: Geneseqp2003bs:*
8: Geneseqp2004s:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	135	100.0	26	AAW53268	AAW53268 Herpes si
2	97	71.9	30	AAW53267	AAW53267 Herpes si
3	54	40.0	129	ABU44880	ABU44880 Protein e
4	54	40.0	129	ABU47370	ABU47370 Protein e
5	50.5	37.4	861	ABU15699	ABU15699 Protein e
6	49.5	36.7	242	ABP07438	ABP07438 Human ORF
7	49.5	36.7	335	AAU59217	AAU59217 Propionib
8	49.5	36.7	335	ABM5736	ABM5736 Propionib
9	49	36.3	871	ADB70269	ADB70269 C. neofo
10	49	36.3	940	ABU35828	ABU35828 Protein e
11	49	36.3	2387	AAU01183	AAU01183 Rat gluta
12	49	36.3	2390	AAU05494	AAU05494 Beta1it s
13	48.5	35.9	526	ABM60918	ABM60918 Drosophil
14	48	35.6	314	ABM60867	ABM60867 Mevalonat
15	48	35.6	314	ABU97197	ABU97197 Enzyme po
16	48	35.6	1404	ADC00560	ADC00560 Enteroha
17	47	34.8	96	AAU90692	AAU90692 Human ple
18	47	34.8	105	ABJ38524	ABJ38524 Human ple
19	47	34.8	151	ABH55486	ABH55486 Lactococc
20	47	34.8	152	AAW49708	AAW49708 Human mon
21	47	34.8	297	AAW21554	AAW21554 Bleomycin
22	47	34.8	333	ABM69430	ABM69430 Photocorin
23	47	34.8	350	ABP65231	ABP65231 Hypoxia-r
24	47	34.8	350	AD558336	AD558336 Human Pro
25	47	34.8	350	AD558334	AD558334 Rat Prote

ALIGNMENTS

25	47	34.8	350	7	AD525739	AD525739 Human pro
27	47	34.8	405	5	ABG70128	ABG70128 Human pre
28	47	34.8	465	5	AAW54351	AAW54351 Vimentin.
29	47	34.8	465	5	ABM77394	ABM77394 Human vim
30	47	34.8	465	7	AD663881	AD663881 Rat Prote
31	47	34.8	466	3	AAU92335	AAU92335 Human vim
32	47	34.8	466	3	AAU92635	AAU92635 Human pol
33	47	34.8	466	4	AAU66348	AAU66348 Human vim
34	47	34.8	466	4	AAU66349	AAU66349 Murine vi
35	47	34.8	466	5	ABM57310	ABM57310 Mouse isc
36	47	34.8	466	5	AAU87694	AAU87694 Human pan
37	47	34.8	466	6	ABR59724	ABR59724 Human vim
38	46	34.1	194	6	ABP73043	ABP73043 Amino aci
39	46	34.1	213	6	ABP73042	ABP73042 Amino aci
40	46	34.1	298	6	ADA34938	ADA34938 Acinetoba
41	46	34.1	387	3	AAU43783	AAU43783 Amino aci
42	46	34.1	387	3	AAU54081	AAU54081 Enzyme EP
43	46	34.1	403	5	ABM57243	ABM57243 Mouse isc
44	46	34.1	496	5	ABM65784	ABM65784 Bifidobac
45	46	34.1	803	6	ABU02124	ABU02124 S. pneumo

RESULT 1

AAW53268 standard; peptide; 26 AA.

AAW53268;
01-JUL-1998 (first entry)
Herpes simplex virus type 1 antiviral agent peptide 2.

Herpes simplex virus type 1; HSV-1; antiviral; inhibition; infection; replication; UL8; POL.

Synthetic.
Herpes simplex virus unknown type.

W09804707-A1.

05-FEB-1998.

28-JUL-1997; 97MO-GB002025.

26-JUL-1996; 96GB-00015730.

(MEDT-) MEDICAL RES COUNCIL.

Marsden HS, Stow ND, Mclean GW;

WPI; 1998-130695/12.

Antiviral agent capable of inhibiting herpes virus replication - disrupts association between herpes simplex virus type 1 UL8 and POL.

Claim 5; Page 57; 83pp; English.

The present sequence represents an antiviral agent peptide capable of combatting herpes virus replication. The antiviral agent is capable of disrupting the association between UL8 and POL (UL30), where UL8 and POL are respectively defined as UL8 and POL of herpes simplex virus type 1 (HSV-1) together with homologues in other herpes viruses. The present invention also describes an assay to determine the ability of a test substance to interfere with the association of UL8 and POL. The assay comprises: (a) exposing a 1st viral component to a test substance followed by a 2nd viral component, or exposing a 1st viral component to a 2nd viral component followed by a test substance; (b) washing to remove any 2nd viral component and/or test substance not associated with the 1st viral component; and (c) detecting the presence, and optionally with the 1st determining the amount, of 2nd viral component associated with the 1st

CC viral component. The antiviral agent can be used to combat herpes virus
 CC replication or infection
 XX
 SQ Sequence 26 AA;

Query Match 100.0%; Score 135; DB 2; Length 26;
 Best Local Similarity 100.0%; Pred. No. 7.3e-13;
 Matches 26; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Cy 1 TWLEERDEWVRSIAVDQAQHAARRVAS 26
 1 TWLEERDEWVRSIAVDQAQHAARRVAS 26

RESULT 2
 AAW53267
 ID AAW53267 standard; peptide; 30 AA.

AAW53267;

01-JUL-1998 (first entry)

Herpes simplex virus type 1 antiviral agent peptide 1.

Herpes simplex virus type 1; HSV-1; antiviral; inhibition; infection;
 replication; UL8; POL.

Synthetic.

Herpes simplex virus unknown type.

WO9804707-A1.

05-FEB-1998.

28-JUL-1997; 97WO-GB002025.

26-JUL-1996; 96GB-00015730.

(MEDI-) MEDICAL RES COUNCIL.

Marsden HS, Stow ND, McLean GW;

WPI; 1998-130695/12.

Antiviral agent capable of inhibiting herpes virus replication - disrupts
 association between herpes simplex virus type 1 UL8 and POL.

Claim 5; Page 57; 83pp; English.

The present sequence represents an antiviral agent peptide capable of
 combating herpes virus replication. The antiviral agent is capable of
 disrupting the association between UL8 and POL (UL30), where UL8 and POL
 are respectively defined as UL8 and POL of herpes simplex virus type 1
 (HSV-1) together with homologues in other herpes viruses. The present
 invention also describes an assay to determine the ability of a test
 substance to interfere with the association of UL8 and POL. The assay
 comprises: (a) exposing a 1st viral component to a test substance
 followed by a 2nd viral component, or exposing a 1st viral component to a
 2nd viral component followed by a test substance; (b) washing to remove
 any 2nd viral component and/or test substance not associated with the 1st
 viral component; and (c) detecting the presence, and optionally
 determining the amount, of 2nd viral component associated with the 1st
 viral component. The antiviral agent can be used to combat herpes virus
 replication or infection

Sequence 30 AA;

Query Match 71.9%; Score 97; DB 2; Length 30;
 Best Local Similarity 95.0%; Pred. No. 3.6e-07;
 Matches 19; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Cy 7 DEWVRSIAVDQAQHAARRVAS 26
 1 DEWVRSIAVDQAQHAARRVAS 26

Db 1 DEWVRSIAVDQAQHAARRVAS 20

RESULT 3

ABU44880
 ID ABU44880 standard; protein; 129 AA.

ABU44880;

19-JUN-2003 (first entry)

Protein encoded by prokaryotic essential gene #30407.

Antisense; prokaryotic essential gene; cell proliferation; drug design.

Salmonella paratyphi.

WO20027183-A2.

03-OCT-2002.

21-MAR-2002; 2002WO-US009107.

21-MAR-2001; 2001US-00815242.

06-SEP-2001; 2001US-00948993.

25-OCT-2001; 2001US-0343923P.

08-FEB-2002; 2002US-00072851.

06-MAR-2002; 2002US-0362699P.

(ELIT-) ELITRA PHARM INC.

Wang L, Zamudio C, Malone C, Haselbeck R, Ohlsen KL, Zyskind JW;

Wall D, Trawick JD, Carr GJ, Yamamoto R, Forsyth RA, Xu HH;

WPI; 2003-029926/02.

N-PSDB; ACA48750.

New antisense nucleic acids, useful for identifying proteins or screening
 for homologous nucleic acids required for cellular proliferation to
 isolate candidate molecules for rational drug discovery programs.

Claim 25; SEQ ID NO 72804; 1766pp; English.

The invention relates to an isolated nucleic acid comprising any one of
 the 6213 antisense sequences given in the specification where expression
 of the nucleic acid inhibits proliferation of a cell. Also included are:
 (1) a vector comprising a promoter operably linked to the nucleic acid
 encoding a polypeptide whose expression is inhibited by the antisense
 nucleic acid; (2) a host cell containing the vector; (3) an isolated
 polypeptide or its fragment whose expression is inhibited by the
 antisense nucleic acid; (4) an antibody capable of specifically binding
 the polypeptide; (5) producing the polypeptide; (6) inhibiting cellular
 proliferation or the activity of a gene in an operon required for
 proliferation; (7) identifying a compound that influences the activity of
 the gene product or that has an activity against a biological pathway
 required for proliferation, or that inhibits cellular proliferation; (8)
 identifying a gene required for cellular proliferation or the biological
 pathway in which a proliferation-required gene or its gene product lies
 or a gene on which the test compound that inhibits proliferation of an
 organism acts; (9) manufacturing an antibiotic; (10) profiling a
 compound's activity; (11) a culture comprising strains in which the gene
 product is overexpressed or underexpressed; (12) determining the extent
 to which each of the strains is present in a culture or collection of
 strains; or (13) identifying the target of a compound that inhibits the
 proliferation of an organism. The antisense nucleic acids are useful for
 identifying proteins or screening for homologous nucleic acids required
 for cellular proliferation to isolate candidate molecules for rational
 drug discovery programs, or for screening homologous nucleic acids
 required for proliferation in cells other than *S. aureus*, *S. typhimurium*,
K. pneumoniae or *P. aeruginosa*. The present sequence is encoded by one of
 the target prokaryotic essential genes. Note: The sequence data for this
 patent did not form part of the printed specification, but was obtained
 in electronic format directly from WIPO at

CC ftp.wipo.int/pub/published_pct_sequences
 XX
 SQ Sequence 129 AA;
 Query Match 40.0%; Score 54; DB 6; Length 129;
 Best Local Similarity 50.0%; Pred. No. 4.2;
 Matches 9; Conservative 3; Mismatches 6; Indels 0; Gaps 0;
 2 WLEERDEWVRSVLAVDAQH 19
 ||||| ||:|:
 37 WREERDSWKGDPLALNGQY 54
 Db
 RESULT 4
 ABU47370
 ID ABU47370 standard; protein; 129 AA.
 XX
 AC ABU47370;
 XX
 DT 19-JUN-2003 (first entry)
 XX
 DE Protein encoded by Prokaryotic essential gene #32897.
 XX
 KM Antisense; prokaryotic essential gene; cell proliferation; drug design.
 XX
 OS Salmomella typhimurium.
 XX
 PN WO200277183-A2.
 XX
 PD 03-OCT-2002.
 XX
 PF 21-MAR-2002; 2002WO-US009107.
 XX
 PR 21-MAR-2001; 2001US-00815242.
 XX
 PR 06-SEP-2001; 2001US-00948993.
 XX
 PR 25-OCT-2001; 2001US-0342923P.
 XX
 PR 08-FEB-2002; 2002US-00072851.
 XX
 PR 06-MAR-2002; 2002US-0362699P.
 XX
 PA (ELIT-) ELITRA PHARM INC.
 XX
 PI Wang L, Zamudio C, Malone C, Haselbeck R, Ohlsen KL, Zyskind JW;
 PI Wall D, Trawick JD, Carr GJ, Yamamoto R, Forsyth RA, Xu HH;
 XX
 DR WPI: 2003-029926/02.
 DR N-PSDB; ACA51240.
 XX
 PT New antisense nucleic acids, useful for identifying proteins or screening
 PT for homologous nucleic acids required for cellular proliferation to
 PT isolate candidate molecules for rational drug discovery programs.
 XX
 PS Claim 25; SEQ ID NO 75294; 1766pp; English.
 XX
 CC The invention relates to an isolated nucleic acid comprising any one of
 CC the 6213 antisense sequences given in the specification where expression
 CC of the nucleic acid inhibits proliferation of a cell. Also included are:
 CC (1) a vector comprising a promoter operably linked to the nucleic acid
 CC encoding a polypeptide whose expression is inhibited by the antisense
 CC nucleic acid; (2) a host cell containing the vector; (3) an isolated
 CC polypeptide or its fragment whose expression is inhibited by the
 CC antisense nucleic acid; (4) an antibody capable of specifically binding
 CC the polypeptide; (5) producing the polypeptide; (6) inhibiting cellular
 CC proliferation or the activity of a gene in an operon required for
 CC proliferation; (7) identifying a compound that influences the activity of
 CC the gene product or that has an activity against a biological pathway
 CC required for proliferation, or that inhibits cellular proliferation; (8)
 CC identifying a gene required for cellular proliferation or the biological
 CC pathway in which a proliferation-required gene or its gene product lies
 CC or a gene on which the test compound that inhibits proliferation of an
 CC organism acts; (9) manufacturing an antibiotic; (10) profiling a
 CC compound's activity; (11) a culture comprising strains in which the gene
 CC product is overexpressed or underexpressed; (12) determining the extent
 CC to which each of the strains is present in a culture or collection of

CC strains; or (13) identifying the target of a compound that inhibits the
 CC proliferation of an organism. The antisense nucleic acids are useful for
 CC identifying proteins or screening for homologous nucleic acids required
 CC for cellular proliferation to isolate candidate molecules for rational
 CC drug discovery programs, or for screening homologous nucleic acids
 CC required for proliferation in cells other than *S. aureus*, *S. typhimurium*,
 CC *K. pneumoniae* or *P. aeruginosa*. The present sequence is encoded by one of
 CC the target prokaryotic essential genes. Note: The sequence data for this
 CC patent did not form part of the printed specification, but was obtained
 CC in electronic format directly from WIPO at
 CC ftp.wipo.int/pub/published_pct_sequences
 XX
 SQ Sequence 129 AA;
 Query Match 40.0%; Score 54; DB 6; Length 129;
 Best Local Similarity 50.0%; Pred. No. 4.2;
 Matches 9; Conservative 3; Mismatches 6; Indels 0; Gaps 0;
 2 WLEERDEWVRSVLAVDAQH 19
 ||||| ||:|:
 37 WREERDSWKGDPLALNGQY 54
 Db
 RESULT 5
 ABU15699
 ID ABU15699 standard; protein; 861 AA.
 XX
 AC ABU15699;
 XX
 DT 19-JUN-2003 (first entry)
 XX
 DE Protein encoded by Prokaryotic essential gene #1226.
 XX
 KM Antisense; prokaryotic essential gene; cell proliferation; drug design.
 XX
 OS *Pseudomonas aeruginosa*.
 XX
 PN WO200277183-A2.
 XX
 PD 03-OCT-2002.
 XX
 PF 21-MAR-2002; 2002WO-US009107.
 XX
 PR 21-MAR-2001; 2001US-00815242.
 XX
 PR 06-SEP-2001; 2001US-00948993.
 XX
 PR 25-OCT-2001; 2001US-0342923P.
 XX
 PR 08-FEB-2002; 2002US-00072851.
 XX
 PR 06-MAR-2002; 2002US-0362699P.
 XX
 PA (ELIT-) ELITRA PHARM INC.
 XX
 PI Wang L, Zamudio C, Malone C, Haselbeck R, Ohlsen KL, Zyskind JW;
 PI Wall D, Trawick JD, Carr GJ, Yamamoto R, Forsyth RA, Xu HH;
 XX
 DR WPI: 2003-029926/02.
 DR N-PSDB; ACA19569.
 XX
 PT New antisense nucleic acids, useful for identifying proteins or screening
 PT for homologous nucleic acids required for cellular proliferation to
 PT isolate candidate molecules for rational drug discovery programs.
 XX
 PS Claim 25; SEQ ID NO 43623; 1766pp; English.
 XX
 CC The invention relates to an isolated nucleic acid comprising any one of
 CC the 6213 antisense sequences given in the specification where expression
 CC of the nucleic acid inhibits proliferation of a cell. Also included are:
 CC (1) a vector comprising a promoter operably linked to the nucleic acid
 CC encoding a polypeptide whose expression is inhibited by the antisense
 CC nucleic acid; (2) a host cell containing the vector; (3) an isolated
 CC polypeptide or its fragment whose expression is inhibited by the
 CC antisense nucleic acid; (4) an antibody capable of specifically binding
 CC the polypeptide; (5) producing the polypeptide; (6) inhibiting cellular
 CC proliferation or the activity of a gene in an operon required for

CC proliferation; (7) identifying a compound that influences the activity of
 CC the gene product or that has an activity against a biological pathway
 CC required for proliferation, or that inhibits cellular proliferation; (8)
 CC identifying a gene required for cellular proliferation or the biological
 CC pathway in which a proliferation-required gene or its gene product lies
 CC or a gene on which the test compound that inhibits proliferation of an
 CC organism acts; (9) manufacturing an antibiotic; (10) profiling a
 CC compound's activity; (11) a culture comprising strains in which the gene
 CC product is overexpressed or underexpressed; (12) determining the extent
 CC to which each of the strains is present in a culture or collection of
 CC strains; or (13) identifying the target of a compound that inhibits the
 CC proliferation of an organism. The antisense nucleic acids are useful for
 CC identifying proteins or screening for homologous nucleic acids required
 CC for cellular proliferation to isolate candidate molecules for rational
 CC drug discovery programs; or for screening homologous nucleic acids
 CC required for proliferation in cells other than *S. aureus*, *S. typhimurium*,
 CC *K. pneumoniae* or *P. aeruginosa*. The present sequence is encoded by one of
 CC the target prokaryotic essential genes. Note: The sequence data for this
 CC patent did not form part of the printed specification, but was obtained
 CC in electronic format directly from WIPO at
 CC ftp.wipo.int/pub/published_pct_sequences

XX Sequence 861 AA;

Query Match 37.4%; Score 50.5; DB 6; Length 861;
 Best Local Similarity 43.5%; Pred. No. 1.1e+02;
 Matches 10; Conservative 3; Mismatches 7; Indels 3; Gaps 1;

QY 2 TWLERDEWV--RLAVDAQHAA 21
 DB 819 WEEDRDMLGRWRXADDPHAA 841

RESULT 6
 ABP07438
 ID ABP07438 standard; protein: 242 AA.
 AC ABP07438;
 XX 25-JUN-2002 (first entry)

XX Human ORFX protein sequence SEQ ID NO:14858.
 DE Human ORFX protein sequence SEQ ID NO:14858.

XX Human; open reading frame; ORFX; gene therapy; cancer; cirrhosis;
 KW hyperproliferative disorder; psoriasis; benign tumour; haemorrhage;
 KW degenerative disorder; osteoarthritis; neurodegenerative disorder;
 KW cardiovascular disease; diabetes mellitus; systemic lupus erythematosus;
 KW hyperextension; hypochyroidism; cholesterol ester storage disease;
 KW immune deficiency; immune disorder; infectious disease;
 KW autoimmune disorder; rheumatoid arthritis; autoimmune thyroiditis;
 KW myasthenia gravis.

XX Homo sapiens.

OS WO200192523-A2.

XX 06-DEC-2001.

PD 29-MAY-2001; 2001WO-US010636.

XX 30-MAY-2000; 2000US-0206132P.

PR 29-AUG-2000; 2000US-0228716P.

XX (CURA-) CURAGEN CORP.

XX Shimkels RA, Leach MD;

XX WPI; 2002-106308/14.

DR N-PSDB; ABN23190.

PT Novel human polypeptides and polynucleotides useful for diagnosing,
 PT preventing and treating cardiovascular disease, neurodegenerative,
 PT hyperproliferative disorders and autoimmune disorders.

XX PS Disclosure; SEQ ID NO 14858; 1037bp; English.

XX The present invention describes substantially purified human proteins
 CC (referred to as open reading frame, ORFX, where X is 1-11491 (see Table 1
 CC in the specification) ABN15762 to ABN22752 encode the human ORFX
 CC proteins given in ABP00010 to ABP11500. ORFX proteins are useful for
 CC treating or preventing a pathology associated with an ORFX-associated
 CC syndrome in humans, and in the manufacture of a medicament for treating a
 CC syndrome associated with ORFX-associated disorder. ORFX polynucleotide
 CC sequences can be used in gene therapy. ORFX sequences can be used in the
 CC treatment of cancer, hyperproliferative disorders, cirrhosis of liver,
 CC psoriasis, benign tumours, keloid, degenerative disorders, haemorrhage,
 CC osteoarthritis, neurodegenerative disorders, disorders related to organ
 CC transplantation, cardiovascular diseases, diabetes mellitus, systemic
 CC lupus erythematosus, hypertension, hypochyroidism, cholesterol ester
 CC storage disease, various immune deficiencies and disorders, infectious
 CC diseases, autoimmune disorders such as multiple sclerosis, rheumatoid
 CC arthritis, autoimmune thyroiditis, myasthenia gravis, graft-versus-host
 CC disease and autoimmune inflammatory eye disease. ORFX proteins are also
 CC useful for treating burns, incisions, ulcers, for treating osteoporosis,
 CC bone degenerative disorders, or periodontal disease, and for gut
 CC protection or regeneration and treatment of lung or liver fibrosis,
 CC reperfusion injury in various tissues and conditions resulting from
 CC systemic cytokine damage. N.B. The sequence data for this patent did not
 CC form part of the printed specification, but was obtained in electronic
 CC format directly from WIPO at ftp.wipo.int/pub/published_pct_sequences

XX Sequence 242 AA;

Query Match 36.7%; Score 49.5; DB 5; Length 242;
 Best Local Similarity 36.1%; Pred. No. 39;
 Matches 13; Conservative 2; Mismatches 8; Indels 13; Gaps 2;

QY 1 TWLERDEWV-----RLAVDAQHAA 24
 DB 162 TWLPRREWVGHSGQALFIEGRGRID-QRVVARY 196

RESULT 7
 AAU59217
 ID AAU59217 standard; protein: 335 AA.

XX AAU59217;

XX 27-FEB-2002 (first entry)

DE Propionibacterium acnes immunogenic protein #20113.

XX SAPHO syndrome; synovitis; acne; pustulosis; hypertosis; osteomyelitis;
 KW uveitis; endophthalmitis; bone; joint; central nervous system; ELISA;
 KW inflammatory lesion; acne vulgaris; enzyme linked immunosorbent assay;
 KW dermatological; osteopathic; neuroprotectant.

XX Propionibacterium acnes.

OS WO200181561-A2.

XX 01-NOV-2001.

PD 20-APR-2001; 2001WO-US012865.

XX 21-APR-2000; 2000US-0199047P.

PR 02-JUN-2000; 2000US-0209841P.

PR 07-JUL-2000; 2000US-0216747P.

XX (CORI-) CORIXA CORP.

XX Skelky YAM, Persing DH, Mitcham UL, Wang SS, Bhatia A;
 PI L maisonneuve J, Zhang Y, Jen S, Carter D;
 XX WPI; 2001-616774/71.

DR N-PSDB; AAS59601.

XX Propionibacterium acnes polypeptides and nucleic acids useful for
PT vaccinating against and diagnosing infections, especially useful for
PT treating acne vulgaris.

XX Example 1, SEQ ID NO 20412; 10699P; English.

CC Sequences AAU39105-AAU68017 represent Propionibacterium acnes immunogenic
CC polypeptides. The proteins and their associated DNA sequences are used in
CC the treatment, prevention and diagnosis of medical conditions caused by
CC P. acnes. The disorders include SAPHO syndrome (synovitis, acne,
CC pustulosis, hyperostosis and osteomyelitis), warts and endophthalmitis.
CC P. acnes is also involved in infections of bone, joints and the central
CC nervous system, however it is particularly involved in the inflammatory
CC lesions associated with acne vulgaris. A method for detecting the
CC presence or absence of P. acnes in a patient comprises contacting a
CC sample with a binding agent that binds to the proteins of the invention
CC and determining the amount of bound protein in the sample. The
CC polypeptides may be used as antigens in the production of antibodies
CC specific for P. acnes proteins. These antibodies can be used to
CC downregulate expression and activity of P. acnes polypeptides and
CC therefore treat P. acnes infections. The antibodies may also be used as
CC diagnostic agents for determining P. acnes presence, for example, by
CC enzyme linked immunosorbent assay (ELISA). Note: The sequence data for
CC this patent did not form part of the printed specification, but was
CC obtained in electronic format directly from WIPO at
CC ftp.wipo.int/pub/published_pct_sequences

XX Sequence 335 AA;

SQ Query Match 36.7%; Score 49.5; DB 4; Length 335;

Best Local Similarity 36.1%; Pred. No. 56;

Matches 13; Conservative 2; Mismatches 8; Indels 13; Gaps 2;

OY 1 TWLEERDEWV-----RSLAVDAQHARRV 24
DB 230 TWLPREEWVGPHSGQALFLGARGRRID-QRVRRV 264

RESULT 8
ABMS5736
ID ABMS5736 standard; protein; 335 AA.

XX ABMS5736;

XX 20-OCT-2003 (first entry)

XX Propionibacterium acnes predicted ORF-encoded polypeptide #20412.

XX Acne vulgaris; antiseborrheic; dermatologically; antibacterial;

XX immunostimulant; immune response; vaccine.

XX Propionibacterium acnes.

XX WO2003033515-A1.

XX 24-APR-2003.

XX 11-OCT-2002; 2002WO-US032727.

XX 15-OCT-2001; 2001US-00978825.

XX (CORI-) CORIXA CORP.

XX Mitcham JL, Skeiky YAW, Persing DH, Bhactia A, Maisonneuve JL;

XX Zhang Y, Wang S, Jen S, Lodes MJ, Benson DR, Jones R, Carter D,

XX Barth B, Valiieve-Douglas J;

XX MPI: 2003-381789/36.

XX N-PSDB; ACR64530.

PT New Propionibacterium acnes polypeptides and polynucleotides encoding the
PT polypeptide, useful for diagnosing, preventing or treating acne vulgaris,

PT or for stimulating an immune response specific for a P. acnes protein.

XX Example 1, SEQ ID NO 20412; 1481P; English.

CC The invention relates to an isolated polynucleotide (ACF64435-ACF64733)
CC encoding a Propionibacterium acnes protein. The invention also relates to
CC polypeptides encoded by the polynucleotides (ABM3624-ABM64536) and to
CC immunogenic fragments of P. acnes polypeptides. The invention
CC additionally encompasses expression vectors and host cells comprising a
CC polynucleotide of the invention; antibodies against polypeptides of the
CC invention; fusion proteins comprising a polypeptide of the invention; a
CC method for stimulating an immune response specific for a P. acnes
CC polypeptide and an isolated T cell population comprising T cells prepared
CC via this method; a vaccine composition (comprising P. acnes polypeptides,
CC polynucleotides, antibodies, fusion proteins, T cell populations, or
CC antigen-presenting cells that express the polypeptide); a method and kit
CC for detecting or determining the presence or absence of P. acnes in a
CC patient; and a method for inhibiting the development of P. acnes in a
CC patient. The P. acnes polypeptides, polynucleotides, antibodies, fusion
CC proteins, T cell populations or antigen-presenting cells that express the
CC polypeptides are useful for diagnosing, preventing or treating acne
CC vulgaris, or for stimulating an immune response specific for a P. acnes
CC protein. The polynucleotides can also be used as probes or primers for the
CC nucleic acid hybridization. The vaccine composition is useful for the
CC stimulation of an immune response against P. acnes, or for treating acne,
CC and the kit is useful for performing a diagnostic assay. The present
CC sequence represents a polypeptide predicted to be encoded by an ORF (open
CC reading frame) contained within the P. acnes polynucleotides of the
CC invention. Note: The sequence data for this patent did not form part of
CC the printed specification, but was obtained in electronic format directly
CC from WIPO at ftp.wipo.int/pub/published_pct_sequences

XX Sequence 335 AA;

SQ Query Match 36.7%; Score 49.5; DB 6; Length 335;

Best Local Similarity 36.1%; Pred. No. 56;

Matches 13; Conservative 2; Mismatches 8; Indels 13; Gaps 2;

OY 1 TWLEERDEWV-----RSLAVDAQHARRV 24
DB 230 TWLPREEWVGPHSGQALFLGARGRRID-QRVRRV 264

RESULT 9
ADB70269
ID ADB70269 standard; protein; 871 AA.

XX ADB70269;

XX 04-DEC-2003 (first entry)

XX C. neoformans amino acid sequence SEQ ID NO:3313.

XX fungicide; gene therapy; infection.

XX Cryptococcus neoformans.

XX WO2003052076-A2.

XX 26-JUN-2003.

XX -17-DEC-2002; 2002WO-US040225.

XX 17-DEC-2001; 2001US-0341261P.

XX (ELIT-) ELITRA PHARM INC.

XX Zamudio C, Eroshtkin AM;

XX MPI: 2003-533017/50.

XX N-PSDB; ADB69186.

PT New nucleic acid, useful for preparing a composition for treating an

Fri Oct 1 12:26:17 2004

us-09-988-851a-12.rag

Page 9

Search completed: September 30, 2004, 11:04:53
Job time : 70.3524 secs

This Page Blank (uspto)

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: September 30, 2004, 11:08:46 ; Search time 75.4286 Seconds
(without alignments)
127.988 Million cell updates/sec

Title: US-09-988-851A-11

Perfect score: 151
Sequence: 1 DEWVRSIAVDQAQAKRVASGIRFFRLNA 30

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 1351062 seqs, 321799191 residues

Total number of hits satisfying chosen parameters: 1351062

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%

Listing first 45 summaries

Database : Published Applications AA:*

1: /cgn2_6/ptodata/1/pubpaa/US07_PUBCOMB.pep:*
2: /cgn2_6/ptodata/1/pubpaa/PT_NEW_PUB.pep:*
3: /cgn2_6/ptodata/1/pubpaa/US06_NEW_PUB.pep:*
4: /cgn2_6/ptodata/1/pubpaa/US06_PUBCOMB.pep:*
5: /cgn2_6/ptodata/1/pubpaa/US07_NEW_PUB.pep:*
6: /cgn2_6/ptodata/1/pubpaa/PTUS_PUBCOMB.pep:*
7: /cgn2_6/ptodata/1/pubpaa/US08_NEW_PUB.pep:*
8: /cgn2_6/ptodata/1/pubpaa/US08_PUBCOMB.pep:*
9: /cgn2_6/ptodata/1/pubpaa/US09_PUBCOMB.pep:*
10: /cgn2_6/ptodata/1/pubpaa/US09B_PUBCOMB.pep:*
11: /cgn2_6/ptodata/1/pubpaa/US09C_PUBCOMB.pep:*
12: /cgn2_6/ptodata/1/pubpaa/US09C_NEW_PUB.pep:*
13: /cgn2_6/ptodata/1/pubpaa/US10A_PUBCOMB.pep:*
14: /cgn2_6/ptodata/1/pubpaa/US10A_PUBCOMB.pep:*
15: /cgn2_6/ptodata/1/pubpaa/US10C_PUBCOMB.pep:*
16: /cgn2_6/ptodata/1/pubpaa/US10C_NEW_PUB.pep:*
17: /cgn2_6/ptodata/1/pubpaa/US60_NEW_PUB.pep:*
18: /cgn2_6/ptodata/1/pubpaa/US60_PUBCOMB.pep:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	151	100.0	30	12	US-09-988-851-11
2	97	64.2	26	12	US-09-988-851-12
3	51	33.8	389	9	US-09-767-770A-2
4	51	33.8	421	9	US-09-767-770A-4
5	51	33.8	421	12	US-10-087-192-540
6	51	33.8	611	10	US-09-893-519A-28
7	50	33.1	347	12	US-10-282-122A-63A55
8	49.5	32.8	77	12	US-10-424-599-203332
9	49.5	32.8	3745	14	US-10-205-032-14
10	49.5	32.8	4471	14	US-10-205-032-10
11	49	32.5	1533	16	US-10-408-765A-869
12	49	32.5	1533	16	US-10-408-765A-870
13	48	31.8	314	14	US-10-166-225A-86
14	48	31.8	421	9	US-09-908-805B-14
15	48	31.8	421	12	US-10-087-192-537

16	48	31.8	445	14	US-10-156-761-9004	Sequence 9004, Ap
17	47.5	31.5	471	15	US-10-369-493-4578	Sequence 4578, Ap
18	47.5	31.5	479	15	US-10-369-493-7335	Sequence 7335, Ap
19	47	31.1	122	16	US-10-767-701-60983	Sequence 60983, A
20	47	31.1	151	16	US-10-767-701-7070	Sequence 37070, A
21	47	31.1	305	12	US-10-424-599-168924	Sequence 168924, A
22	47	31.1	315	14	US-10-156-761-9370	Sequence 9370, Ap
23	47	31.1	821	16	US-10-437-963-132870	Sequence 132870, Ap
24	47	31.1	831	16	US-10-437-963-132836	Sequence 132836, A
25	47	31.1	1459	16	US-10-437-963-162463	Sequence 162463, A
26	46.5	30.8	226	12	US-10-425-114-65804	Sequence 65804, A
27	46.5	30.8	784	12	US-10-156-761-11603	Sequence 11603, A
28	46	30.5	100	16	US-10-437-963-137743	Sequence 137743, A
29	46	30.5	172	16	US-10-437-963-118015	Sequence 118015, A
30	46	30.5	175	16	US-10-767-701-53805	Sequence 53805, A
31	46	30.5	195	14	US-10-307-138-23	Sequence 23, Appl
32	46	30.5	275	12	US-10-425-114-45133	Sequence 45133, A
33	46	30.5	312	14	US-10-156-761-7723	Sequence 7723, Ap
34	46	30.5	399	15	US-10-369-493-9142	Sequence 9142, Ap
35	46	30.5	436	16	US-10-388-566-1288	Sequence 1288, Ap
36	46	30.5	440	16	US-10-389-566-1580	Sequence 1580, Ap
37	46	30.5	505	16	US-10-437-963-136760	Sequence 136760, Ap
38	46	30.5	5588	16	US-10-329-148A-6	Sequence 6, Appl
39	45.5	30.1	251	10	US-09-880-748-1411	Sequence 1411, Ap
40	45.5	30.1	251	12	US-10-293-418-1411	Sequence 1411, Ap
41	45.5	30.1	502	15	US-10-369-493-18430	Sequence 18430, A
42	45.5	30.1	502	16	US-10-388-566-1132	Sequence 1132, Ap
43	45.5	30.1	502	16	US-10-389-566-2076	Sequence 2076, Ap
44	45	29.8	79	16	US-10-437-963-178266	Sequence 178266, Ap
45	45	29.8	113	12	US-10-282-122A-59280	Sequence 59280, A

ALIGNMENTS

RESULT 1
US-09-988-851-11
Sequence 11, Application US/0998851
Publication NO. US20020034609A1
GENERAL INFORMATION:
APPLICANT: Medical Research Council
APPLICANT: Marsden, Howard
APPLICANT: Stow, Nigel
APPLICANT: McLean, Gordon
TITLE OF INVENTION: ASSAY FOR SCREENING FOR AN ANTI-VIRAL AGENT
FILE REFERENCE: P29347A
CURRENT APPLICATION NUMBER: US/09/998, 851
CURRENT FILING DATE: 2002-03-11
NUMBER OF SEQ ID NOS: 18
SOFTWARE: PatentIn version 3.1
SEQ ID NO 11
LENGTH: 30
TYPE: PRT
ORGANISM: Artificial Sequence
FEATURE:
OTHER INFORMATION: Peptide derived from herpes simplex virus.
US-09-988-851-11

Query Match 100.0% Score 151; DB 12; Length 30;
Best Local Similarity 100.0% Pred No. 3.9e-16;
Matches 30; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 DEWVRSIAVDQAQAKRVASGIRFFRLNA 30
Db 1 DEWVRSIAVDQAQAKRVASGIRFFRLNA 30

RESULT 2
US-09-988-851-12
Sequence 12, Application US/0998851
Publication NO. US20020034609A1
GENERAL INFORMATION:
APPLICANT: Medical Research Council

APPLICANT: Marsden, Howard
APPLICANT: Stow, Nigel
TITLE OF INVENTION: ASSAY FOR SCREENING FOR AN ANTI-VIRAL AGENT
FILE REFERENCE: P29347A
CURRENT APPLICATION NUMBER: US/09/998,851
CURRENT FILING DATE: 2002-03-11
NUMBER OF SEQ ID NOS: 18
SOFTWARE: PatentIn version 3.1
SEQ ID NO 12
LENGTH: 26
TYPE: PRT
ORGANISM: Artificial Sequence
FEATURE:
OTHER INFORMATION: Peptide derived from herpes simplex virus.
US-09-998-851-12

Query Match 64.2%; Score 97; DB 12; Length 26;
Best Local Similarity 95.0%; Pred. No. 8.3e-08;
Matches 19; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy 1 DEWVSLAVDQAHAKRVAS 20
Db 7 DEWVSLAVDQAHAKRVAS 26

RESULT 3
US-09-767-770A-2
Sequence 2, Application US/09767770A
Patent No. US20020065237A1
GENERAL INFORMATION:
APPLICANT: Michalovich, David
APPLICANT: Sims, Matthew Alan
TITLE OF INVENTION: NOVEL COMPOUNDS
FILE REFERENCE: GP-30086-C1
CURRENT APPLICATION NUMBER: US/09/767,770A
CURRENT FILING DATE: 2001-01-23
PRIOR APPLICATION NUMBER: 9806222.7
PRIOR FILING DATE: 1998-03-23
PRIOR APPLICATION NUMBER: 9820299.7
PRIOR FILING DATE: 1998-09-17
PRIOR APPLICATION NUMBER: 09/266,261
PRIOR FILING DATE: 1999-03-11
NUMBER OF SEQ ID NOS: 4
SOFTWARE: FastSeq for Windows Version 3.0
SEQ ID NO 2
LENGTH: 389
TYPE: PRT
ORGANISM: HOMO SAPIENS
US-09-767-770A-2

Query Match 33.8%; Score 51; DB 9; Length 389;
Best Local Similarity 35.5%; Pred. No. 24;
Matches 11; Conservative 7; Mismatches 11; Indels 2; Gaps 1;

Qy 1 DEWVSLAVDQAHAKRVAS 29
Db 311 DRWVSVSFSDHGLHVASLADDKVRFWRID 341

RESULT 4
US-09-767-770A-4
Sequence 4, Application US/09767770A
Patent No. US20020065237A1
GENERAL INFORMATION:
APPLICANT: Michalovich, David
APPLICANT: Sims, Matthew Alan
TITLE OF INVENTION: NOVEL COMPOUNDS
FILE REFERENCE: GP-30086-C1
CURRENT APPLICATION NUMBER: US/09/767,770A
CURRENT FILING DATE: 2001-01-23

PRIOR APPLICATION NUMBER: 9806222.7
PRIOR FILING DATE: 1998-03-23
PRIOR APPLICATION NUMBER: 9820299.7
PRIOR FILING DATE: 1998-09-17
PRIOR APPLICATION NUMBER: 09/266,261
PRIOR FILING DATE: 1999-03-11
NUMBER OF SEQ ID NOS: 4
SOFTWARE: FastSeq for Windows Version 3.0
SEQ ID NO 4
LENGTH: 421
TYPE: PRT
ORGANISM: HOMO SAPIENS
US-09-767-770A-4

Query Match 33.8%; Score 51; DB 9; Length 421;
Best Local Similarity 35.5%; Pred. No. 26;
Matches 11; Conservative 7; Mismatches 11; Indels 2; Gaps 1;

Qy 1 DEWVSLAVDQAHAKRVAS 29
Db 311 DRWVSVSFSDHGLHVASLADDKVRFWRID 341

RESULT 5
US-10-087-192-540
Sequence 540, Application US/10087192
Publication No. US20020182586A1
GENERAL INFORMATION:
APPLICANT: Morris, David W.
APPLICANT: Engelhard, Eric K.
TITLE OF INVENTION: NOVEL COMPOSITIONS AND METHODS FOR
FILE REFERENCE: 529452000122
CURRENT APPLICATION NUMBER: US/10/087,192
CURRENT FILING DATE: 2002-03-01
PRIOR APPLICATION NUMBER: US 09/747,377
PRIOR FILING DATE: 2000-12-22
PRIOR APPLICATION NUMBER: US 09/798,586
PRIOR FILING DATE: 2001-03-02
NUMBER OF SEQ ID NOS: 2059
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 540
LENGTH: 421
TYPE: PRT
ORGANISM: Homo sapiens
US-10-087-192-540

Query Match 33.8%; Score 51; DB 12; Length 421;
Best Local Similarity 35.5%; Pred. No. 26;
Matches 11; Conservative 7; Mismatches 11; Indels 2; Gaps 1;

Qy 1 DEWVSLAVDQAHAKRVAS 29
Db 311 DRWVSVSFSDHGLHVASLADDKVRFWRID 341

RESULT 6
US-09-893-519A-28
Sequence 28, Application US/09893519A
Publication No. US20030027243A1
GENERAL INFORMATION:
APPLICANT: ANADYS PHARMACEUTICALS, INC.
APPLICANT: THOMPSON, Craig
APPLICANT: MOORE, Jeffrey
APPLICANT: BURMAN, Ed T.
APPLICANT: BRADLEY, John
APPLICANT: DESILVA, Thamara
APPLICANT: HARRIS, Sandra
APPLICANT: KOMARNITSKY, Svetlana
APPLICANT: MENDILLO, Marc
APPLICANT: MOORE, Daniel
APPLICANT: MCCOY, Melissa
APPLICANT: SANDERSON, Karen


```
APPLICANT: HAO, Tairig
APPLICANT: ZHU, Shuhao
APPLICANT: LONG, Fan
APPLICANT: DAVIDOV, Eugene
TITLE OF INVENTION: ANTI-FUNGAL COMPOUNDS AND METHODS OF USE
FILE REFERENCE: 0342/1G548-US2
CURRENT APPLICATION NUMBER: US/09/893,519A
CURRENT FILING DATE: 2001-06-28
PRIOR APPLICATION NUMBER: US 60/225,164
PRIOR FILING DATE: 2000-06-29
PRIOR APPLICATION NUMBER: US 60/224,457
PRIOR FILING DATE: 2000-08-10
NUMBER OF SEQ ID NOS: 146
SOFTWARE: PatentIn version 3.1
SEQ ID NO 28
LENGTH: 611
TYPE: PRT
ORGANISM: Homo sapiens
FEATURE:
NAME/KEY: misc feature
OTHER INFORMATION: Corresponds to SEQ ID NO: 101
PUBLICATION INFORMATION:
DATABASE ACCESSION NUMBER: Human Genbank/NP_001078
DATABASE ENTRY DATE: 2001-12-18
RELEVANT RESIDUES: (1)..(611)
US-09-893-519A-28

Query Match      33.8%; Score 51; DB 10; Length 611;
Best Local Similarity 35.5%; Pred. No. 39;
Matches 11; Conservative 7; Mismatches 11; Indels 2; Gaps 1;

QY      1 DEWVSLAV--DAQHAKRVASEGLRFFRLN 29
Db      501 DRWVSVSFSHGHVAVSLADKVRFRWRID 531

RESULT 7
US-10-282-122A-63456
; Sequence 63456, Application US/10282122A
; Publication No. US20040029129A1
; GENERAL INFORMATION:
; APPLICANT: Wang, Liangsu
; APPLICANT: Zamudio, Carlos
; APPLICANT: Malone, Cheryl
; APPLICANT: Haselbeck, Robert
; APPLICANT: Ohlsen, Kari
; APPLICANT: Zyskind, Judith
; APPLICANT: Wall, Daniel
; APPLICANT: Trawick, John
; APPLICANT: Carr, Grant
; APPLICANT: Yamamoto, Robert
; APPLICANT: Forsyth, R.
; APPLICANT: Xu, H.
TITLE OF INVENTION: Identification of Essential Genes in Microorganisms
FILE REFERENCE: ELITRA.034A
CURRENT APPLICATION NUMBER: US/10/282,122A
CURRENT FILING DATE: 2003-02-20
PRIOR APPLICATION NUMBER: 60/191,078
PRIOR FILING DATE: 2000-03-21
PRIOR APPLICATION NUMBER: 60/206,848
PRIOR FILING DATE: 2000-05-23
PRIOR APPLICATION NUMBER: 60/207,727
PRIOR FILING DATE: 2000-05-26
PRIOR APPLICATION NUMBER: 60/230,335
PRIOR FILING DATE: 2000-09-06
PRIOR APPLICATION NUMBER: 60/230,347
PRIOR FILING DATE: 2000-09-09
PRIOR APPLICATION NUMBER: 60/242,578
PRIOR FILING DATE: 2000-10-23
PRIOR APPLICATION NUMBER: 60/253,625
PRIOR FILING DATE: 2000-11-27
PRIOR APPLICATION NUMBER: 60/257,931
PRIOR FILING DATE: 2000-12-22
```

```

; PRIOR APPLICATION NUMBER: 60/267,636
; PRIOR FILING DATE: 2001-02-09
; PRIOR APPLICATION NUMBER: 60/269,308
; PRIOR FILING DATE: 2001-02-16
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 78614
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 63456
; LENGTH: 347
; TYPE: PRT
; ORGANISM: Mycoplasma genitalium
US-10-282-122A-63456

Query Match      33.1%; Score 50; DB 12; Length 347;
Best Local Similarity 42.3%; Pred. No. 30;
Matches 11; Conservative 5; Mismatches 10; Indels 0; Gaps 0;

QY      1 DEWVSLAVDAQHAKRVASEGLRFF 26
Db      309 DEWVSLKLVNDGKQAKVADETLKM 334

RESULT 8
US-10-424-599-203332
; Sequence 203332, Application US/10424599
; Publication No. US20040031072A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa, Thomas J
; APPLICANT: Kovalic, David K
; APPLICANT: Zhou, Yihua
; APPLICANT: Cao, Yongwei
TITLE OF INVENTION: Soy Nucleic Acid Molecules and Other Molecules Associated With
FILE REFERENCE: 38-21(53223)B
CURRENT APPLICATION NUMBER: US/10/424,599
CURRENT FILING DATE: 2003-04-28
NUMBER OF SEQ ID NOS: 285684
SEQ ID NO 203332
LENGTH: 77
TYPE: PRT
ORGANISM: Glycine max
FEATURE:
OTHER INFORMATION: Clone ID: PAT_MPT3847_25633C.1.pep
US-10-424-599-203332

Query Match      32.8%; Score 49.5; DB 12; Length 77;
Best Local Similarity 35.3%; Pred. No. 6.7;
Matches 12; Conservative 8; Mismatches 7; Indels 7; Gaps 1;

QY      4 VRSLA-----VDAQHAKRVASEGLRFFRLNA 30
Db      19 IRLSLADLNKRTIIQSTYSKASQATRFYIINA 52

RESULT 9
US-10-205-032-14
; Sequence 14, Application US/10205032
; Publication No. US20030113874A1
; GENERAL INFORMATION:
; APPLICANT: Farnet, Chris
; APPLICANT: Yang, Xianshu
; APPLICANT: Scatfa, Alfredo
TITLE OF INVENTION: GENES AND PROTEINS FOR THE BIOSYNTHESIS OF ROSARAMICIN
FILE REFERENCE: 3016-2US
CURRENT APPLICATION NUMBER: US/10/205,032
CURRENT FILING DATE: 2002-07-26
NUMBER OF SEQ ID NOS: 39
SOFTWARE: PatentIn version 3.0
SEQ ID NO 14
LENGTH: 3745
TYPE: PRT
ORGANISM: microsporidia carbonacea subspecies aurantiaca
US-10-205-032-14
```

Query Match 32.8%; Score 49.5; DB 14; Length 3745;
 Best Local Similarity 54.5%; Pred. No. 5.1e+02;
 Matches 12; Conservative 3; Mismatches 6; Indels 1; Gaps 1;

QY 3 WVRSLAVD-AQHAARKVASEGL 23
 DB 781 WVRRAVDYASHSAHVAVEGM 802

RESULT 10

US-10-205-032-10
 ; Sequence 10, Application US/10205032
 ; Publication No. US20030113874A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Farnet, Chris
 ; APPLICANT: Yang, Xianhu
 ; APPLICANT: Statia, Alitico
 ; TITLE OF INVENTION: GENES AND PROTEINS FOR THE BIOSYNTHESIS OF ROSAPAMICIN
 ; FILE REFERENCE: 30.6-2US
 ; CURRENT APPLICATION NUMBER: US/10/205,032
 ; CURRENT FILING DATE: 2002-07-26
 ; NUMBER OF SEQ ID NOS: 39
 ; SOFTWARE: PatentIn version 3.0
 ; SEQ ID NO 10
 ; LENGTH: 4471
 ; TYPE: PRT
 ; ORGANISM: micromonospora carbonacea subspecies aurantiaca
 US-10-205-032-10

Query Match 32.8%; Score 49.5; DB 14; Length 4471;
 Best Local Similarity 54.5%; Pred. No. 6.2e+02;
 Matches 12; Conservative 3; Mismatches 6; Indels 1; Gaps 1;

QY 3 WVRSLAVD-AQHAARKVASEGL 23
 DB 752 WVRRAVDYASHSAHVAVEGM 773

RESULT 11

US-10-408-765A-869
 ; Sequence 869, Application US/10408765A
 ; Publication No. US20040101874A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Ghosh, Soumitra S.
 ; APPLICANT: Fany, Boia D.
 ; APPLICANT: Zhang, Bing
 ; APPLICANT: Gibson, Bradford W.
 ; APPLICANT: Taylor, Steven W.
 ; APPLICANT: Glenn, Gary M.
 ; APPLICANT: Warnock, Dale E.
 ; TITLE OF INVENTION: TARGETS FOR THERAPEUTIC INTERVENTION
 ; TITLE OF INVENTION: IDENTIFIED IN THE MITOCHONDRIAL PROTEOME
 ; FILE REFERENCE: 660088.465
 ; CURRENT APPLICATION NUMBER: US/10/408,765A
 ; CURRENT FILING DATE: 2003-04-04
 ; NUMBER OF SEQ ID NOS: 3077
 ; SOFTWARE: FastSeq for Windows Version 4.0
 ; SEQ ID NO 869
 ; LENGTH: 1533
 ; TYPE: PRT
 ; ORGANISM: Homo sapiens
 US-10-408-765A-869

Query Match 32.5%; Score 49; DB 16; Length 1533;
 Best Local Similarity 31.0%; Pred. No. 2.2e+02;
 Matches 9; Conservative 7; Mismatches 13; Indels 0; Gaps 0;

QY 1 DEWVSLAVDAQHAARKVASEGLRFFRLN 29
 DB 526 DEWIKTISAEIQDELSRTDYEQKRPDQKN 554

RESULT 12

US-10-408-765A-870
 ; Sequence 870, Application US/10408765A
 ; Publication No. US20040101874A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Ghosh, Soumitra S.
 ; APPLICANT: Fany, Boia D.
 ; APPLICANT: Zhang, Bing
 ; APPLICANT: Gibson, Bradford W.
 ; APPLICANT: Taylor, Steven W.
 ; APPLICANT: Glenn, Gary M.
 ; APPLICANT: Warnock, Dale E.
 ; TITLE OF INVENTION: TARGETS FOR THERAPEUTIC INTERVENTION
 ; TITLE OF INVENTION: IDENTIFIED IN THE MITOCHONDRIAL PROTEOME
 ; FILE REFERENCE: 660088.465
 ; CURRENT APPLICATION NUMBER: US/10/408,765A
 ; CURRENT FILING DATE: 2003-04-04
 ; NUMBER OF SEQ ID NOS: 3077
 ; SOFTWARE: FastSeq for Windows Version 4.0
 ; SEQ ID NO 870
 ; LENGTH: 1533
 ; TYPE: PRT
 ; ORGANISM: Homo sapiens
 US-10-408-765A-870

Query Match 32.5%; Score 49; DB 16; Length 1533;
 Best Local Similarity 31.0%; Pred. No. 2.2e+02;
 Matches 9; Conservative 7; Mismatches 13; Indels 0; Gaps 0;

QY 1 DEWVSLAVDAQHAARKVASEGLRFFRLN 29
 DB 526 DEWIKTISAEIQDELSRTDYEQKRPDQKN 554

RESULT 13

US-10-166-225A-86
 ; Sequence 86, Application US/10166225A
 ; Publication No. US20030148416A1
 ; GENERAL INFORMATION:
 ; APPLICANT: BERRY, Alan
 ; APPLICANT: BRETZEL, Werner
 ; APPLICANT: HUMBELIN, Markus
 ; APPLICANT: LOPEZ-ULIBARRI, Rual
 ; APPLICANT: MAYER, Anne P.
 ; APPLICANT: YELISEEV, Alexei A.
 ; TITLE OF INVENTION: IMPROVED ISOPRENOID PRODUCTION
 ; FILE REFERENCE: C38435/121966
 ; CURRENT APPLICATION NUMBER: US/10/166,225A
 ; CURRENT FILING DATE: 2002-06-05
 ; NUMBER OF SEQ ID NOS: 197
 ; SOFTWARE: PatentIn version 3.1
 ; SEQ ID NO 86
 ; LENGTH: 314
 ; TYPE: PRT
 ; ORGANISM: Streptococcus pyogenes
 US-10-166-225A-86

Query Match 31.8%; Score 48; DB 14; Length 314;
 Best Local Similarity 61.5%; Pred. No. 55;
 Matches 8; Conservative 1; Mismatches 4; Indels 0; Gaps 0;

QY 1 DEWVSLAVDAQH 13
 DB 200 DEWVQSAIDYQH 212

RESULT 14

US-09-908-805B-14
 ; Sequence 14, Application US/09908805B
 ; Patent No. US20020147307A1
 ; GENERAL INFORMATION:
 ; APPLICANT: HILTON, Douglas J
 ; APPLICANT: ALEXANDER, Warren S

; APPLICANT: VINEY, Elizabeth M
 ; APPLICANT: WILSON, Tracey A
 ; APPLICANT: RICHARDSON, Rachael T
 ; APPLICANT: STARR, Robyn
 ; APPLICANT: NICHOLSON, Sandra E
 ; APPLICANT: METCALF, Donald
 ; APPLICANT: NICOLA, Nicos A
 ; TITLE OF INVENTION: THERAPEUTIC AND DIAGNOSTIC AGENTS
 ; FILE REFERENCE: 109762
 ; CURRENT APPLICATION NUMBER: US/09/908,805B
 ; CURRENT FILING DATE: 2001-07-19
 ; PRIOR APPLICATION NUMBER: 08/962,560
 ; PRIOR FILING DATE: 1997-10-31
 ; NUMBER OF SEQ ID NOS: 81
 ; SOFTWARE: PatentIn Ver. 2.0
 ; SEQ ID NO 14
 ; LENGTH: 421
 ; TYPE: PRT
 ; ORGANISM: Mouse
 ; US-09-908-805B-14

Query Match 31.8%; Score 48; DB 9; Length 421;
 Best Local Similarity 35.5%; Pred. No. 76;
 Matches 11; Conservative 6; Mismatches 12; Indels 2; Gaps 1;

QY 1 DEWVR--SLAVDAQHAKRVASEGLRFFRLN 29
 Db 311 DRWVRAVSFSDHGLVSLADDKWRFWRID 341

RESULT 15
 ; US-10-087-192-537
 ; Sequence 537, Application US/10087192
 ; Publication No. US20020182586A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Morris, David W.
 ; APPLICANT: Engelhard, Eric K.
 ; TITLE OF INVENTION: NOVEL COMPOSITIONS AND METHODS FOR
 ; TITILE OF INVENTION: CANCER
 ; FILE REFERENCE: 529452000122
 ; CURRENT APPLICATION NUMBER: US/10/087,192
 ; CURRENT FILING DATE: 2002-03-01
 ; PRIOR APPLICATION NUMBER: US 09/747,377
 ; PRIOR FILING DATE: 2000-12-22
 ; PRIOR APPLICATION NUMBER: US 09/798,586
 ; PRIOR FILING DATE: 2001-01-02
 ; NUMBER OF SEQ ID NOS: 2059
 ; SOFTWARE: FastSeq for Windows Version 4.0
 ; SEQ ID NO 537
 ; LENGTH: 421
 ; TYPE: PRT
 ; ORGANISM: Mus musculus
 ; US-10-087-192-537

Query Match 31.8%; Score 48; DB 12; Length 421;
 Best Local Similarity 35.5%; Pred. No. 76;
 Matches 11; Conservative 6; Mismatches 12; Indels 2; Gaps 1;

QY 1 DEWVR--SLAVDAQHAKRVASEGLRFFRLN 29
 Db 311 DRWVRAVSFSDHGLVSLADDKWRFWRID 341

Search completed: September 30, 2004, 11:33:35
 Job time : 76.4286 secs

This Page Blank (uspto)

Fri Oct 1 12:26:15 2004

us-09-988-851a-11.ra1

Page 1

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OW protein - protein search, using sw model

Run on: September 30, 2004, 10:58:43 ; Search time 20.8571 Seconds
(without alignments)
74.257 Million cell updates/sec

Title: US-09-988-851a-11
Perfect score: 151
Sequence: 1 DEWVRSIAYDAQHAKRVASEGIRFFRLNA 30

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 389414 seqs, 51625971 residues

Total number of hits satisfying chosen parameters: 389414

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Issued Patents AA:*
1: /cgn2_6/ptodata/2/1aa/5A-COMB.pep.*
2: /cgn2_6/ptodata/2/1aa/5B-COMB.pep.*
3: /cgn2_6/ptodata/2/1aa/6A-COMB.pep.*
4: /cgn2_6/ptodata/2/1aa/6B-COMB.pep.*
5: /cgn2_6/ptodata/2/1aa/PCITUS-COMB.pep.*
6: /cgn2_6/ptodata/2/1aa/backfile1.pep.*

Pred. NO. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	151	100.0	30	4	US-09-230-405-11
2	97	64.2	26	4	US-09-230-405-12
3	50	33.1	233	4	US-09-328-352-4522
4	49.5	32.8	919	4	US-09-543-681A-7187
5	49.5	32.5	484	4	US-09-252-991A-16729
6	48	31.8	421	4	US-09-303-765-14
7	48	31.8	1871	2	US-08-694-869-1
8	48	31.8	1871	2	US-09-349-546-1
9	48	31.8	1871	4	US-09-502-831-1
10	47	31.1	742	4	US-09-252-991A-32049
11	46	30.5	212	4	US-09-540-236-2005
12	46	30.5	404	4	US-09-252-991A-19166
13	46	30.5	829	3	US-09-413-814-105
14	46	30.5	5588	3	US-09-036-967A-6
15	46	30.5	5588	3	US-09-370-700-6
16	46	30.5	5588	4	US-09-603-207-6
17	45	29.8	209	4	US-09-252-991A-28592
18	45	29.8	261	4	US-09-489-039A-11649
19	45	29.8	339	4	US-09-252-991A-26745
20	45	29.8	547	4	US-09-252-991A-25705
21	44.5	28.5	557	4	US-09-134-000C-4941
22	44	29.1	274	4	US-09-489-039A-14140
23	44	29.1	401	4	US-09-462-645C-2
24	44	29.1	401	4	US-09-462-645C-6
25	44	29.1	401	4	US-09-462-645C-10
26	44	29.1	574	2	US-08-756-317-4
27	44	29.1	600	2	US-08-756-317-3

28	44	29.1	638	2	US-08-756-317-2	Sequence 2, Appli
29	44	29.1	925	4	US-09-540-236-3586	Sequence 3586, Ap
30	44	29.1	984	4	US-09-716-865-12	Sequence 12, Appl
31	44	29.1	7257	3	US-09-335-403-5	Sequence 5, Appli
32	44	29.1	7257	4	US-09-568-102-5	Sequence 5, Appli
33	44	29.1	7257	4	US-09-567-968-5	Sequence 5, Appli
34	44	29.1	7257	4	US-09-568-488-5	Sequence 5, Appli
35	44	29.1	7257	4	US-09-568-486-5	Sequence 5, Appli
36	44	29.1	7257	4	US-09-568-472-5	Sequence 5, Appli
37	44	29.1	7257	4	US-09-567-899-5	Sequence 5, Appli
38	43	28.5	259	4	US-09-252-991A-27735	Sequence 27735, A
39	43	28.5	302	4	US-09-252-991A-21528	Sequence 21528, A
40	43	28.5	339	3	US-09-120-365-62	Sequence 62, Appl
41	43	28.5	339	3	US-09-120-365-80	Sequence 80, Appl
42	43	28.5	339	3	US-09-515-039-62	Sequence 62, Appl
43	43	28.5	339	3	US-09-515-039-80	Sequence 80, Appl
44	43	28.5	661	4	US-09-252-991A-29083	Sequence 29083, A
45	43	28.5	742	4	US-09-489-039A-9962	Sequence 9962, Ap

ALIGNMENTS

RESULT 1
US-09-230-405-11
Sequence 11, Application US/09230405
Patent No. 6337074
GENERAL INFORMATION:
APPLICANT: Medical Research Council
TITLE OF INVENTION: ANTI-VIRAL AGENT
FILE REFERENCE: PI7970C
CURRENT FILING DATE: 1999-01-25
CURRENT FILING DATE: 1999-01-25
NUMBER OF SEQ ID NOS: 18
SOFTWARE: Patent In Ver. 2.1
SEQ ID NO 11
LENGTH: 30
TYPE: PRT
ORGANISM: Artificial Sequence
FEATURE:
OTHER INFORMATION: Description of Artificial Sequence:peptide derived
OTHER INFORMATION: from herpes simplex virus
US-09-230-405-11

Query Match 100.0% Score 151; DB 4; Length 30;
Best Local Similarity 100.0% Pred. No. 3.2e-18;
Matches 30; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Cy 1 DEWVRSIAYDAQHAKRVASEGIRFFRLNA 30
Db 1 DEWVRSIAYDAQHAKRVASEGIRFFRLNA 30
RESULT 2
US-09-230-405-12
Sequence 12, Application US/09230405
Patent No. 6337074
GENERAL INFORMATION:
APPLICANT: Medical Research Council
TITLE OF INVENTION: ANTI-VIRAL AGENT
FILE REFERENCE: PI7970C
CURRENT FILING DATE: 1999-01-25
CURRENT FILING DATE: 1999-01-25
NUMBER OF SEQ ID NOS: 18
SOFTWARE: Patent In Ver. 2.1
SEQ ID NO 12
LENGTH: 26
TYPE: PRT
ORGANISM: Artificial Sequence
FEATURE:
OTHER INFORMATION: Description of Artificial Sequence:peptide derived
OTHER INFORMATION: from herpes simplex virus
US-09-230-405-12

Query Match 64.2%; Score 97; DB 4; Length 26;
Best Local Similarity 95.0%; Pred. No. 2.4e-09;
Matches 19; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 DEWVSLAVDAQHAKRVAS 20
DB 7 DEWVSLAVDAQHAKRVAS 26

RESULT 3
US-09-328-352-4522
Sequence 4522, Application US/09328352
Patent No. 6562958
GENERAL INFORMATION:
APPLICANT: Gary L. Breton et al.
TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO ACINETOBACTER
FILE REFERENCE: GTC99-033PA
CURRENT APPLICATION NUMBER: US/09/328.352
CURRENT FILING DATE: 1999-06-04
NUMBER OF SEQ ID NOS: 8252
SEQ ID NO 4522
LENGTH: 233
TYPE: PRT
ORGANISM: Acinetobacter baumannii
US-09-328-352-4522

Query Match 33.1%; Score 50; DB 4; Length 233;
Best Local Similarity 60.0%; Pred. No. 2.2;
Matches 12; Conservative 2; Mismatches 6; Indels 0; Gaps 0;

QY 5 RSLAVDAQHAKRVASEGLR 24
DB 77 RAKVDAQTLAKRAAEGR 96

RESULT 4
US-09-543-681A-7187
Sequence 7187, Application US/09543681A
Patent No. 6605709
GENERAL INFORMATION:
APPLICANT: GARY BRETON
TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PROTEUS MIRABILIS
FILE REFERENCE: 2709.1002-001
CURRENT APPLICATION NUMBER: US/09/543.681A
CURRENT FILING DATE: 2000-04-05
PRIOR APPLICATION NUMBER: US 60/128,706
PRIOR FILING DATE: 1999-04-09
NUMBER OF SEQ ID NOS: 8344
SEQ ID NO 7187
LENGTH: 919
TYPE: PRT
ORGANISM: Proteus mirabilis
US-09-543-681A-7187

Query Match 32.8%; Score 49.5; DB 4; Length 919;
Best Local Similarity 43.3%; Pred. No. 15;
Matches 13; Conservative 5; Mismatches 9; Indels 3; Gaps 1;

QY 1 DEWVSLAVDAQHAKRVASEGLRFFRLNA 30
DB 736 NEWLSLPEKASPOVKRVAT---RFAMLDA 762

RESULT 5
US-09-252-991A-16729
Sequence 16729, Application US/09252991A
Patent No. 6551795
GENERAL INFORMATION:
APPLICANT: Marc J. Rubenfield et al.
TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS

TITLE OF INVENTION: AERUGINOSA FOR DIAGNOSTICS AND THERAPEUTICS
FILE REFERENCE: 107196.136
CURRENT APPLICATION NUMBER: US/09/252.991A
CURRENT FILING DATE: 1999-02-18
PRIOR APPLICATION NUMBER: US 60/074,788
PRIOR FILING DATE: 1998-02-18
PRIOR APPLICATION NUMBER: US 60/094,130
PRIOR FILING DATE: 1998-07-27
NUMBER OF SEQ ID NOS: 33142
SEQ ID NO 16729
LENGTH: 484
TYPE: PRT
ORGANISM: Pseudomonas aeruginosa
US-09-252-991A-16729

Query Match 32.5%; Score 49; DB 4; Length 484;
Best Local Similarity 50.0%; Pred. No. 8;
Matches 13; Conservative 3; Mismatches 8; Indels 2; Gaps 1;

QY 7 LAVDAQHAKRVASE--GLRFFRLNA 30
DB 273 LAADRAHARRVPAQHLGLAGFRIGA 298

RESULT 6
US-09-302-769-14
Sequence 14, Application US/09302769
Patent No. 6323317
GENERAL INFORMATION:
APPLICANT: HILTON, Douglas J
APPLICANT: ALEXANDER, Warren S
APPLICANT: VINEY, Elizabeth M
APPLICANT: WILSON, Tracey A
APPLICANT: RICHARDSON, Rachel T
APPLICANT: STAR, Robyn
APPLICANT: NICHOLSON, Sandra E
APPLICANT: METCALF, Donald
APPLICANT: NICOLA, Nicos A
TITLE OF INVENTION: THERAPEUTIC AND DIAGNOSTIC AGENTS
FILE REFERENCE: 109762
CURRENT APPLICATION NUMBER: US/09/302.769
CURRENT FILING DATE: 1999-04-30
PRIOR APPLICATION NUMBER: 08/962,560
PRIOR FILING DATE: 1997-10-31
SOFTWARE: PatentIn Ver. 2.0
NUMBER OF SEQ ID NOS: 50
SEQ ID NO 14
LENGTH: 421
TYPE: PRT
ORGANISM: Mouse
US-09-302-769-14

Query Match 31.8%; Score 48; DB 4; Length 421;
Best Local Similarity 35.5%; Pred. No. 9.9;
Matches 11; Conservative 6; Mismatches 12; Indels 2; Gaps 1;

QY 1 DEWV--SLAVDAQHAKRVASEGLRFFRLN 29
DB 311 DRWVAVSFSDGLHVASLADDKVRFWRID 341

RESULT 7
US-08-694-869-1
Sequence 1, Application US/08694869
Patent No. 5994123
GENERAL INFORMATION:
APPLICANT: Olszewski, N.
APPLICANT: Tsafir, I.
APPLICANT: Somers, D. A.
APPLICANT: Lockhart, B.
APPLICANT: Torbert, K.
TITLE OF INVENTION: SUGARCANE BACILLIFORM VIRUS
TITLE OF INVENTION: PROMOTER

Fri Oct 1 12:26:15 2004

us-09-988-851a-11.ra1

Page 3

NUMBER OF SEQUENCES: 10
CORRESPONDENCE ADDRESS:
ADDRESSEE: Schwegman, Lundberg, Woessner & Kluth, P.A.
STREET: P.O. Box 2938
CITY: Minneapolis
STATE: MN
COUNTRY: USA
ZIP: 55402
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette
COMPUTER: IBM Compatible
OPERATING SYSTEM: DOS
SOFTWARE: FASTSEQ Version 2.0
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/694,869
FILING DATE: 09-AUG-1996
CLASSIFICATION: 800
PRIOR APPLICATION DATA:
APPLICATION NUMBER:
FILING DATE:
ATTORNEY/AGENT INFORMATION:
NAME: Woessner, Warren D
REGISTRATION NUMBER: 30,440
REFERENCE/DOCKET NUMBER: 600,369US1
TELECOMMUNICATION INFORMATION:
TELEPHONE: 612-373-6900
TELEFAX: 612-339-3061
TELEX:
INFORMATION FOR SEQ ID NO: 1:
SEQUENCE CHARACTERISTICS:
LENGTH: 1871 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-694-869-1

Query Match 31.8%; Score 48; DB 2; Length 1871;
Best Local Similarity 39.1%; Pred. No. 62;
Matches 9; Conservative 6; Mismatches 8; Indels 0; Gaps 0;
CY 8 AVDAQHAKRVASEGRLPFLRNA 30
DB 30 AQHAKHQAQRIAGQALAFRLVTS 52
RESULT 8
US-09-349-546-1
; Sequence 1, Application US/09349546
; Patent No. 6093569
; GENERAL INFORMATION:
; APPLICANT: Olszewski, N.
; APPLICANT: Tzafir, I.
; APPLICANT: Somers, D.A.
; APPLICANT: Lockhart, B.
; APPLICANT: Torbert, K.
; TITLE OF INVENTION: Sugarcane bacilliform virus promoter
; FILE REFERENCE: 600,369US2
; CURRENT APPLICATION NUMBER: US/09/349,546
; EARLIER FILING DATE: 1999-07-08
; EARLIER APPLICATION NUMBER: US 08/694,869
; EARLIER FILING DATE: 1996-08-09
; EARLIER APPLICATION NUMBER: PCT/IB97/01338
; EARLIER FILING DATE: 1997-08-13
; NUMBER OF SEQ ID NOS: 10
; SOFTWARE: FASTSEQ for Windows Version 4.0
; SEQ ID NO: 1
; LENGTH: 1871
; TYPE: PRT
; ORGANISM: sugarcane bacilliform virus
US-09-349-546-1
Query Match 31.8%; Score 48; DB 3; Length 1871;

Best Local Similarity 39.1%; Pred. No. 62;
Matches 9; Conservative 6; Mismatches 8; Indels 0; Gaps 0;
CY 8 AVDAQHAKRVASEGRLPFLRNA 30
DB 30 AQHAKHQAQRIAGQALAFRLVTS 52
RESULT 9
US-09-502-831-1
; Sequence 1, Application US/09502831
; Patent No. 6489462
; GENERAL INFORMATION:
; APPLICANT: Olszewski, Neil
; APPLICANT: Tzafir, Itis
; APPLICANT: Somers, David A.
; APPLICANT: Lockhart, Benham
; APPLICANT: Torbert, Kimberly A.
; TITLE OF INVENTION: Sugarcane Bacilliform Virus Promoter
; FILE REFERENCE: 600,369US3
; CURRENT APPLICATION NUMBER: US/09/502,831
; CURRENT FILING DATE: 2000-02-11
; PRIOR APPLICATION NUMBER: PCT/IB97/01338
; PRIOR FILING DATE: 1997-08-13
; PRIOR APPLICATION NUMBER: US 08/694,869
; PRIOR FILING DATE: 1996-08-09
; NUMBER OF SEQ ID NOS: 10
; SOFTWARE: FASTSEQ for Windows Version 4.0
; SEQ ID NO: 1
; LENGTH: 1871
; TYPE: PRT
; ORGANISM: Sugarcane bacilliform virus
US-09-502-831-1

Query Match 31.8%; Score 48; DB 4; Length 1871;
Best Local Similarity 39.1%; Pred. No. 62;
Matches 9; Conservative 6; Mismatches 8; Indels 0; Gaps 0;
CY 8 AVDAQHAKRVASEGRLPFLRNA 30
DB 30 AQHAKHQAQRIAGQALAFRLVTS 52

RESULT 10
US-09-252-991A-32049
; Sequence 32049, Application US/09252991A
; Patent No. 6551795
; GENERAL INFORMATION:
; APPLICANT: Marc J. Rubenfield et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
; FILE REFERENCE: 107196,136
; CURRENT APPLICATION NUMBER: US/09/252,991A
; CURRENT FILING DATE: 1999-02-18
; PRIOR APPLICATION NUMBER: US 60/074,788
; PRIOR FILING DATE: 1998-02-18
; PRIOR APPLICATION NUMBER: US 60/094,190
; PRIOR FILING DATE: 1998-07-27
; NUMBER OF SEQ ID NOS: 33142
; SEQ ID NO 32049
; LENGTH: 742
; TYPE: PRT
; ORGANISM: Pseudomonas aeruginosa
US-09-252-991A-32049

Query Match 31.1%; Score 47; DB 4; Length 742;
Best Local Similarity 50.0%; Pred. No. 29;
Matches 11; Conservative 4; Mismatches 7; Indels 0; Gaps 0;
CY 1 DEWRSLANDVDAQHAKRVASEG 22
DB 147 DRRRSIALAAQQAERRVVAAG 168

```

RESULT 11
US-09-540-236-2005
; Sequence 2005, Application US/09540236
; Patent No. 6673910
; GENERAL INFORMATION:
; APPLICANT: Gary L. Bretton et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO MORAXELLA CATAR
; FILE REFERENCE: 2709.2005-001
; CURRENT APPLICATION NUMBER: US/09/540,236
; CURRENT FILING DATE: 2000-04-04
; NUMBER OF SEQ ID NOS: 3840
; SEQ ID NO 2005
; LENGTH: 212
; TYPE: PRT
; ORGANISM: M.cattarrhalis
US-09-540-236-2005

Query Match      30.5%; Score 46; DB 4; Length 212;
Best Local Similarity 47.8%; Pred. No. 9.1;
Matches 11; Conservative 2; Mismatches 10; Indels 0; Gaps 0;

QY 1 DEWYRSIAVDAQHAKRVASGLRFF 23
DB 12 DEFWQLTAQSNEMAYHPASVGL 34

RESULT 12
US-09-252-991A-19166
; Sequence 19166, Application US/09252991A
; Patent No. 6551795
; GENERAL INFORMATION:
; APPLICANT: Marc J. Rubenfield et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
; FILE REFERENCE: 107196.136
; CURRENT APPLICATION NUMBER: US/09/252,991A
; CURRENT FILING DATE: 1999-02-18
; PRIOR APPLICATION NUMBER: US 60/074,788
; PRIOR FILING DATE: 1998-02-18
; PRIOR APPLICATION NUMBER: US 60/094,190
; PRIOR FILING DATE: 1998-07-27
; NUMBER OF SEQ ID NOS: 33142
; SEQ ID NO 19166
; LENGTH: 404
; TYPE: PRT
; ORGANISM: Pseudomonas aeruginosa
US-09-252-991A-19166

Query Match      30.5%; Score 46; DB 4; Length 404;
Best Local Similarity 39.3%; Pred. No. 20;
Matches 11; Conservative 4; Mismatches 13; Indels 0; Gaps 0;

QY 1 DEWYRSIAVDAQHAKRVASGLRFF 28
DB 336 DEWYRSIAVDAQHAKRVASGLRFF 363

RESULT 13
US-09-413-814-105
; Sequence 105, Application US/09413814
; Patent No. 6225064
; GENERAL INFORMATION:
; APPLICANT: Gesellschaft fuer Biotechnologische Forschung mbH
; APPLICANT: Bristol-Myers Squibb, Co.
; APPLICANT: Beyer, Stefan
; APPLICANT: Bloeker, Helmut
; APPLICANT: Brandt, Petra
; APPLICANT: Cino, Paul M
; APPLICANT: Dougherty, Brian A
; APPLICANT: Goldberg, Steven L
; APPLICANT: Hofle, Gerhard

```

```

; APPLICANT: Mueller, Joachim
; APPLICANT: Reichenbach, Hans
; TITLE OF INVENTION: DNA sequences for enzymatic synthesis of polyketide or
; TITLE OF INVENTION: heteropolyketide compounds
; FILE REFERENCE: PCT/US 99/23535
; CURRENT APPLICATION NUMBER: US/09/413,814
; CURRENT FILING DATE: 1999-10-07
; EARLIER APPLICATION NUMBER: DE 198 46 493.2
; EARLIER FILING DATE: 1998-10-09
; NUMBER OF SEQ ID NOS: 107
; SOFTWARE: Patent In Ver. 2.1
; SEQ ID NO 105
; LENGTH: 829
; TYPE: PRT
; ORGANISM: Sorangium cellulosum
US-09-413-814-105

Query Match      30.5%; Score 46; DB 3; Length 829;
Best Local Similarity 38.5%; Pred. No. 49;
Matches 10; Conservative 5; Mismatches 9; Indels 2; Gaps 1;

QY 3 WYRSIAVDAQHAKRVASGLRFF 26
DB 579 WYRSIAVDAQHAKRVASGLRFF 604

RESULT 14
US-09-036-987A-6
; Sequence 6, Application US/09036987A
; Patent No. 6143526
; GENERAL INFORMATION:
; APPLICANT: Baitz, Richard H.
; APPLICANT: Broughton, Mary C.
; APPLICANT: Crawford, Kathryn P.
; APPLICANT: Madduri, Krishnamurthy
; APPLICANT: Merlo, Donald J.
; APPLICANT: Tredey, Patil J.
; APPLICANT: Turner, Jan R.
; TITLE OF INVENTION: Biosynthetic Genes For Spinosyn Insecticide
; TITLE OF INVENTION: Production
; NUMBER OF SEQUENCES: 39
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Dow Agrosciences LLC Patent Department
; STREET: 9330 Zionsville Road
; CITY: Indianapolis
; STATE: Indiana
; COUNTRY: USA
; ZIP: 46268
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/036,987A
; FILING DATE: 09-MAR-1998
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Stuart, Donald R.
; REGISTRATION NUMBER: 28,479
; REFERENCE/DOCKET NUMBER: 50,608
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (317)337-4816
; TELEFAX: (317)337-4847
; INFORMATION FOR SEQ ID NO: 6:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 5588 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-09-036-987A-6

```


Query Match 30.5%; Score 46; DB 3; Length 5588;
Best Local Similarity 39.3%; Pred. No. 5.1e+02;
Matches 11; Conservative 4; Mismatches 11; Indels 2; Gaps 1;

QY 1 DEWVRSI--AVDAQHAKRVASEGLRFF 26
DB 799 DYWYONLRQPVRFQQTVARMAPDQGYRFF 826

RESULT 15

US-09-370-700-6
; Sequence 6, Application US/09370700
; Patent No. 6274350
; GENERAL INFORMATION:
; APPLICANT: Baltz, Richard H
; APPLICANT: Broughton, Mary C
; APPLICANT: Crawford, Kathryn P
; APPLICANT: Madduri, Krishnamurthy
; APPLICANT: Treadway, Patti J
; APPLICANT: Turner, Jan R
; APPLICANT: Waldron, Clive
; TITLE OF INVENTION: Biosynthetic Genes For Spinosyn Insecticide
; FILE REFERENCE: 50489 DIV1
; CURRENT APPLICATION NUMBER: US/09/370,700
; EARLIER FILING DATE: 1999-08-09
; EARLIER APPLICATION NUMBER: US 09/36987
; NUMBER OF SEQ ID NOS: 39
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 6
; LENGTH: 5588
; TYPE: PRT
; ORGANISM: Saccharopolyspora spinosa
US-09-370-700-6

Query Match 30.5%; Score 46; DB 3; Length 5588;
Best Local Similarity 39.3%; Pred. No. 5.1e+02;
Matches 11; Conservative 4; Mismatches 11; Indels 2; Gaps 1;

QY 1 DEWVRSI--AVDAQHAKRVASEGLRFF 26
DB 799 DYWYONLRQPVRFQQTVARMAPDQGYRFF 826

Search completed: September 30, 2004, 11:00:06
Job time : 21.8571 secs

This Page Blank (uspto)

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: September 30, 2004, 10:58:43; Search time 11.1429 Seconds
(without alignments)
140.189 Million cell updates/sec

Title: US-09-988-851a-11

Perfect score: 151

Sequence: 1 DEMVSLAVDAQHAKRVASGLRFRFLNA 30

Scoring table: BLOSUM62

Gapop 10.0, Gapext 0.5

Searched: 141681 seqs, 52070155 residues

Total number of hits satisfying chosen parameters: 141681

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-Processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database: SwissProt_42:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	151	100.0	798	HEPA_HCMVA	P16827 human cytom
2	53	35.1	1318	1 VIVD_BPT7	P03726 bacterioph
3	51	33.8	421	WSBI_HUMAN	O96177 homo sapien
4	50	33.1	347	SYW_MYCE	P47372 mycoplasma
5	49	32.5	1533	Y586_HUMAN	O9HYV6 homo sapien
6	48	31.8	342	COX2_BACPF	O04441 bacillus ps
7	48	31.8	421	WSBI_MOUSE	O54927 mus musculu
8	48	31.8	585	SACL_CHLRE	O39593 chlamydomon
9	47.5	31.5	619	ABS_DROME	O93350 drosophila
10	47	31.1	259	UBIE_AGRIS	O81115 agrobacteri
11	47	31.1	354	FAH2_STRAW	O81185 xanthomonas
12	47	31.1	354	DCUP_XANAC	O81185 xanthomonas
13	47	31.1	612	YK66_SCHPO	O91512 schistosom
14	46	30.5	436	1 SECY_METUA	O60175 methanococ
15	46	30.5	568	1 DONS_DROME	O9YNA8 drosophila
16	46	30.5	1242	1 RPO2_ASFB7	P42487 atirican swi
17	45.5	30.1	502	1 KPYK_IACLA	O07637 lactococcus
18	45	29.8	255	1 TRMD_ECOLI	O87410 escherichia
19	45	29.8	255	1 TRMD_SALTI	O87410 escherichia
20	45	29.8	255	1 TRMD_SALTY	P16245 salmonella
21	45	29.8	493	1 SECY_ARCFU	O28377 archaeglob
22	45	29.8	873	1 GLND_VIBRT	O9AC07 vibrio fisc
23	45	29.8	873	1 D33A_MYCE	O49405 mycoplasma
24	45	29.5	265	1 UL07_HSVSA	O01028 herpesvitu
25	44.5	29.5	493	1 AMYR_DROOR	O77015 drosophila
26	44.5	29.5	940	1 SEAI_MYCLR	P57996 mycobacteri
27	44.5	29.5	969	1 YEAC_SCHPO	O14077 schistosac
28	44.5	29.5	266	1 PTNG_ECOLI	P08187 escherichia
29	44	29.1	293	1 Y844_ARCFU	O28414 archaeglob
30	44	29.1	293	1 NUSG_STRVG	P27309 streptomyc
31	44	29.1	299	1 PAH4_STRCO	O926V4 streptomyc
32	44	29.1	316	1 ECHI_RAT	O62651 ratius norv
33	44	29.1	327	1 ECHI_RAT	O62651 ratius norv

ALIGNMENTS

34	44	29.1	333	1	DARF_CIOAB	O97FV2 clostridium
35	44	29.1	354	1	DCUP_XANCP	O606X1 xanthomonas
36	44	29.1	382	1	BDR_BACHD	O9XA69 bacillus ha
37	44	29.1	383	1	BODG_PSEBK	P80193 pseudomonas
38	44	29.1	401	1	CSDA_ECOLI	O46925 escherichia
39	44	29.1	404	1	METR_CHLRE	O8KE97 chlorobium
40	44	29.1	624	1	MUTL_XANCP	O8P864 xanthomonas
41	44	29.1	625	1	MUTL_XANAC	O8P1X2 xanthomonas
42	44	29.1	591	1	VPJ3_YEAST	P20795 saccharomyc
43	43.5	28.8	207	1	Y196_RALSO	P88634 raietonia s
44	43.5	28.8	293	1	TRUB_STRMU	O84U15 streptococ
45	43.5	28.8	329	1	TRUB_IACLA	O9AC00 lactococcus

AC	HEPA_HCMVA	STANDARD;	PRT;	798 AA.
AC	P16827;			
DT	01-AUG-1990 (Rel. 15, Created)			
DT	01-AUG-1990 (Rel. 15, Last sequence update)			
DT	01-OCT-1996 (Rel. 34, Last annotation update)			
DE	DNA helicase/primase complex associated protein.			
GN	UL102.			
OS	Human cytomegalovirus (strain AD169).			
OC	Viruses; dsDNA viruses, no RNA stage; Herpesviridae;			
OC	Betaherpesvirinae; Cytomegalovirus.			
OX	NCBI_TaxID=10360;			
RN	[1]			
RP	SEQUENCE FROM N.A.			
RX	MEDLINE=90269039; PubMed=2161319; Bohni R., Brown C.M., Cerny R.,			
RA	Chee M.S., Bankier A.T., Beck S., Hirschfeld T., Hutchison C.A. III, Kozakaris T., Martignetti J.A.,			
RA	Horstmann E., Satchell S.C., Tomlinson P., Weston K.M., Barrell B.G.;			
RT	"Analysis of the protein-coding content of the sequence of human			
RT	cytomegalovirus strain AD169."			
RL	Chrr. Top. Microbiol. Immunol. 154:125-169(1990).			
CC	- FUNCTION: PROBABLY INVOLVED IN DNA REPLICATION.			
CC	- SIMILARITY: BELONGS TO FAMILY THAT GROUPS TOGETHER HSV-1 UL8,			
CC	EHV-1 54, VZV 52 AND HCMV 102.			
CC	-----			
CC	This SWISS-PROT entry is copyright. It is produced through a collaboration			
CC	between the Swiss Institute of Bioinformatics and the EMBL outstation -			
CC	the European Bioinformatics Institute. There are no restrictions on its			
CC	use by non-profit institutions as long as its content is in no way			
CC	modified and this statement is not removed. Usage by and for commercial			
CC	entities requires a license agreement (see http://www.isb-sib.ch/announce/			
CC	or send an email to license@isb-sib.ch).			
CC	-----			
DR	EMBL; X17403; CAA35338.1; -			
DR	PIR; S03867; S09667.			
KW	DNA replication.			
SO	SEQUENCE 798 AA; 85613 MW; 6A4AE14B1E1F5B CRC64;			

Qy	1	DEMVSLAVDAQHAKRVASGLRFRFLNA 30		
Db	769	DEMVSLAVDAQHAKRVASGLRFRFLNA 798		

Query Match	100.0%; Score 151; DB 1; Length 798;
Best Local Similarity	100.0%; Pred. No. 5.2e-15;
Matches	30; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

RESULT 2	STANDARD;	PRT;	1316 AA.
ID	VIVD_BPT7		
ID	VIVD_BPT7		
AC	P03726;		
DT	21-JUL-1986 (Rel. 01, Created)		
DT	21-JUL-1986 (Rel. 01, Last sequence update)		
DT	10-OCT-2003 (Rel. 42, Last annotation update)		
DE	Internal virion protein D.		

```

GN 16.
OS Bacteriophage T7.
OC Viruses; dsDNA viruses, no RNA stage; Caudovirales; Podoviridae;
OC T7-like viruses.
OX NCBI_TaxId=10760;
RN (1)
RP SEQUENCE FROM N.A.
RX MEDLINE=83241725; PubMed=6864790;
RA Dunn J.J., Studier F.W.;
RT "Complete nucleotide sequence of bacteriophage T7 DNA and the
RT locations of 17 genetic elements.";
RL J. Mol. Biol. 166:477-535(1983).
RN (2)
RP SIMILARITY TO SLT.
RX MEDLINE=94262160; PubMed=8203016;
RA Koonin E.V., Rudd K.E.;
RT "A conserved domain in putative bacterial and bacteriophage
RT transglycosylases.";
RL Trends Biochem. Sci. 19:106-107(1994).
CC -1- FUNCTION: MAY BE INVOLVED IN THE LYSIS OF THE BACTERIAL CELL
CC -1- SIMILARITY: Belongs to the transglycosylase slt family.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL: V01146; CA24434.1; -.
DR PIR: A04352; HIBPD7.
DR InterPro: IPR008258; SLT_dom.
DR InterPro: IPR00189; Transglyc_AS.
DR Pfam: PF01464; SLT; 1.
DR PROSITE: PS00922; TRANSGLYCOSYLASE; 1.
DR Cell wall; Hydrolase; Glycosidase.
FT DOMAIN 24 111 SLT-TYPE DOMAIN.
FT ACT SITE 37 37 BY SIMILARITY.
SQ SEQUENCE 1318 AA; 143838 MW; 51A0AA920CBE210 CRC64;

Query Match 35.1%; Score 53; DB 1; Length 118;
Best Local Similarity 60.0%; Pred. No. 6.9;
Matches 12; Conservative 3; Mismatches 5; Indels 0; Gaps 0;

CY 5 RSLAVDAQHAKRYVASEGLR 24
DB 358 KALVGAESALINVASEGRL 377

RESULT 3
MSBL_HUMAN STANDARD; PRT; 421 AA.
AC Q9Y6I7; Q9NRE1; Q9CBH9; Q9UNN6;
DT 28-FEB-2003 (Rel. 41, Created)
DT 28-FEB-2003 (Rel. 41, Last sequence update)
DT 10-OCT-2003 (Rel. 42, Last annotation update)
DE WD repeat and SOCS box containing protein 1 (MSB-1) (SOCS box-
DE containing WD protein SMIP-1).
GN MSB1 OR SMIP1.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
OX NCBI_TaxId=9606;
OX [1]
RP SEQUENCE FROM N.A. (ISOFORM 1).
RX MEDLINE=99284520; PubMed=10354473;
RA Vasilioukas D., Hancock S., Stern C.D.;
RT "SMIP-1: novel SOCS box containing WD-protein regulated by signalling
RT centres and by Shh during development.";
RL Mech. Dev. 82:79-94(1999).
RN [2]

```

```

RP SEQUENCE FROM N.A. (ISOFORMS 1 AND 2).
RA Zhang J.W., Liu T.X., Shen Y., Chen S.J., Chen Z.;
RL Submitted (MAR-2000) to the EMBL/GenBank/DBJ databases.
RN [3]
RP SEQUENCE FROM N.A. (ISOFORM 1).
RC TISSUE=Brain;
RX MEDLINE=21154917; PubMed=11230166;
RA Wiemann S., Weill B., Wellenreuther R., Gassenhuber J., Glassl S.,
RA Lauber J., Boecker M., Blocker A., Baurasachs S., Blum H.,
RA Lauber U., Duesterhoeft A., Beyer A., Koehler K., Strack N.,
RA Mewes H.-W., Othenwaelder B., Obermaler B., Tampe J., Heubner D.,
RA Wambut R., Korn B., Klein M., Poustka A.;
RT "Towards a catalog of human genes and proteins: sequencing and
RT analysis of 500 novel complete protein coding human cDNAs.";
RL Genome Res. 11:422-435(2001).
RN [4]
RP SEQUENCE FROM N.A. (ISOFORM 1).
RC TISSUE=Hypothalamus;
RX MEDLINE=20402571; PubMed=10931946;
RA Hu R.-M., Han Z.-G., Song H.-D., Peng Y.-D., Huang Q.-H., Ren S.-X.,
RA Gu Y.-J., Huang C.-H., Li Y.-B., Jiang C.-L., Fu G., Zhang Q.-H.,
RA Gu B.-W., Dai M., Mao Y.-F., Gao G.-F., Rong R., Ye M., Zhou J.,
RA Xu S.-H., Gu J., Shi J.-X., Jin W.-R., Zhang C.-K., Wu T.-M.,
RA Huang G.-Y., Chen Z., Chen M.-D., Chen J.-L.;
RT "Gene expression profiling in the human hypothalamus-pituitary-adrenal
RT axis and full-length cDNA cloning.";
RL Proc. Natl. Acad. Sci. U.S.A. 97:9543-9548(2000).
RN [5]
RP SEQUENCE FROM N.A. (ISOFORM 1).
RC TISSUE=pituitary;
RA Song H., Peng Y., Dai M., Huang Q., Mao Y., Zhang Q., Mao M., Fu G.,
RA Luo M., Chen J., Hu R.;
RL Submitted (NOV-1998) to the EMBL/GenBank/DBJ databases.
RN [6]
RP SEQUENCE FROM N.A. (ISOFORM 1).
RC TISSUE=Colon;
RX MEDLINE=22388257; PubMed=12477932;
RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bat N.K.,
RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
RA Datchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
RA Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
RA Brownstein M.U., Ustin T.B., Toshynski S., Carrinpi P., Prange C.,
RA Raha S.S., Loggiano N.A., Peters G.J., Abramson R.D., Mullahy S.J.,
RA Bosak S.A., McMan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Huliy S.W.,
RA Villalón D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
RA Fahy J., Helton E., Kettelman M., Madan A., Rodriguez S., Sanchez A.,
RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,
RA Butcherfield Y.S.N., Krzywinski M.I., Skalska U., Smalins D.E.,
RA Schnerch A., Schein J.E., Jones S.J.M., Marra M.A.;
RT "Generation and initial analysis of more than 15,000 full-length
RT human and mouse cDNA sequences.";
RL Proc. Natl. Acad. Sci. U.S.A. 99:16699-16903(2002).
RN [7]
RP ALTERNATIVE PRODUCTS:
DE Event=Alternative splicing; Named isoforms=2;
Name=1;
Name=2;
Name=2;
Isoid=Q9Y6I7-2; Sequence=VSP 006792;
-1- SIMILARITY: Contains 6 WD repeats.
-1- SIMILARITY: Contains 1 SOCS box domain.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).

```

```

CC -----
DR EMBL: AF072880; AAD28808.1; -
DR EMBL: AF069313; AAD20564.2; -
DR EMBL: AF240696; AAE82746.1; -
DR EMBL: AL110243; CAB53693.1; -
DR EMBL: AF112205; AAF17193.1; -
DR EMBL: AF106683; AAD43036.1; -
DR EMBL: BC021110; AAR21110.1; -
DR PIR: T14773; T14772.
DR InterPro: IPR001496; SOCS.C.
DR Pfam: PF00400; WD40.
DR PRINTS: PR00320; GPOTEINRPT.
DR ProDom: PD000018; WD40; 2.
DR SMART: SM00253; SOCS; 1.
DR SMART: SM00320; WD40; 6.
DR PROSITE: PS00678; WD_REPEATS_1; 2.
DR PROSITE: PS50082; WD_REPEATS_2; 6.
DR PROSITE: PS50294; WD_REPEATS_REGION; 1.
DR Repeat: WD repeat; Alternative splicing.
KW Repeat: WD repeat; Alternative splicing.
FT REPEAT 32 71 WD 1.
FT REPEAT 124 165 WD 2.
FT REPEAT 168 208 WD 3.
FT REPEAT 212 251 WD 4.
FT REPEAT 254 293 WD 5.
FT REPEAT 309 346 WD 6.
FT DOMAIN 372 421 SOCS BOX.
FT VARSPLIC 14 159 Missing (in isoform 2).
FT CONFLICT 16 16 L -> S (IN REF. 3 AND 4).
FT CONFLICT 178 178 F -> L (IN REF. 5).
FT CONFLICT 231 231 S -> P (IN REF. 5).
SQ SEQUENCE 421 AA; 47432 MW; 650D4942E97D4BED CRC64;

Query Match 33.8%; Score 51; DB 1; Length 421;
Best Local Similarity 35.5%; Pred. No. 4.2;
Matches 11; Conservative 7; Mismatches 11; Indels 2; Gaps 1;

Oy 1 DEWVSLAV--DAQAAKRVASEGLRPFRLN 29
Db 311 DRWVSVSFSDHGLHVSADKKVRFWRID 341

RESULT 4
SYM_MYCGE STANDARD; PRT; 347 AA.
ID _SYM_MYCGE
AC P47372;
DT 01-FEB-1996 (Rel. 33, Created)
DT 01-FEB-1996 (Rel. 33, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Tryptophanyl-tRNA synthetase (EC 6.1.1.2) (Tryptophan-tRNA ligase)
DE (TrpRS).
DE TRPS OR MG126.
OS Mycoplasma genitalium.
OC Bacteriia; Firmicutes; Mollicutes; Mycoplasmataceae; Mycoplasma.
OC NCBI_TaxId=2057;
OX NCBI_TaxId=2057;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=ATCC 33530 / G-37;
RX MEDLINE=96026346; PubMed=7569993;
RA Fleisemann R.D., Bult C.J., Kerlavage A.R., Sutton G., Kelley J.M.,
RA Fritschman J.L., Weidman J.F., Small K.V., Sandusky M., Fuhrmann J.L.,
RA Nguyen D.T., Ueberbach T.R., Sauder D.M., Phillips C.A., Merrick J.M.,
RA Tomb J.-F., Dougherty B.A., Bott K.F., Hu P.-C., Lucier T.S.,
RA Peterson S.N., Smith H.O., Hutchison C.A. III, Venter J.C.,
RA "The minimal gene complement of Mycoplasma genitalium.";
RT Science 270:397-403 (1995).
RL Science 270:397-403 (1995).
CC -1- CATALYTIC ACTIVITY: ATP + L-tryptophan + tRNA(Trp) = AMP +
CC dihydrophosphate + L-tryptophanyl-tRNA(Trp).
CC -1- SUBUNIT: Homodimer (by similarity).
CC -1- SUBCELLULAR LOCATION: Cytoplasmic.

```

```

CC -----
CC -1- SIMILARITY: Belongs to class-I aminoacyl-tRNA synthetase family.
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (see http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC EMBL: U39693; AAC71344.1; -
DR PIR: I64213; I64213.
DR HSSP: P00953; 1D2R.
DR TIGR: MG126; -.
DR HAMAP: MF_00140; -.
DR InterPro: IPR002305; tRNA-synt_1b.
DR InterPro: IPR001412; tRNA-synt_1.
DR InterPro: IPR002306; Trp-tRNA-synt_1b.
DR Pfam: PF00579; tRNA-synt_1b; 1.
DR PRINTS: PR01039; TRNASYNTHRP.
DR TIGRPFAM: TIGR00293; trps; 1.
DR PROSITE: PS00178; AA-tRNA_LIGASE_1; FALSE NEG.
DR Aminoacyl-tRNA synthetase; Protein biosynthesis; ligase; ATP-binding;
KW Complete proteome.
FT SITE 11 19 "HIGH" REGION.
FT SITE 200 204 "WMSKS" REGION.
FT BINDING 203 203 ATP (BY SIMILARITY).
SQ SEQUENCE 347 AA; 39328 MW; 82DIC572A523132B CRC64;

Query Match 33.1%; Score 50; DB 1; Length 347;
Best Local Similarity 42.3%; Pred. No. 4.9;
Matches 11; Conservative 5; Mismatches 10; Indels 0; Gaps 0;

Oy 1 DEWVSLAVDAQAAKRVASEGLRPF 26
Db 309 DEWVSLVNDGRKQAKVADETLKM 334

RESULT 5
SYM_HUMAN STANDARD; PRT; 1533 AA.
ID _SYM_HUMAN
AC Q9BYV6; O60328;
DT 28-FEB-2003 (Rel. 41, Created)
DT 28-FEB-2003 (Rel. 41, Last sequence update)
DT 10-OCT-2003 (Rel. 42, Last annotation update)
DE Hypothetical protein KIA0586.
DE KIA0586.
OS Homo sapiens (human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
OC NCBI_TaxId=9606;
OX NCBI_TaxId=9606;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Brain;
RX MEDLINE=98290545; PubMed=9628581;
RA Nagase T., Ishikawa K.-I., Miyajima N., Tanaka A., Kotani H.,
RA Nemura N., Ohara O.;
RA "Prediction of the coding sequences of unidentified human genes. IX.
RA The complete sequences of 100 new cDNA clones from brain which can
RA code for large proteins in vitro.";
RL DNA Res. 5:31-38 (1998).
RL DNA Res. 5:31-38 (1998).
RN [2]
RP SEQUENCE OF 1187-1533 FROM N.A.
RC TISSUE=Placenta;
RX MEDLINE=22388257; PubMed=12477932;
RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
RA Hopkins R.F., Jordan H., Moore T., Max S.L., Wang J., Hsieh F.,
RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
RA Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
RA Brownstein M.J., Ueidi T.B., Toshitoki S., Carninci P., Prange C.,
RA Raha S.S., Loguclano N.A., Peters G.J., Abramson R.D., Mullanly S.J.,

```

```

CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
-----
CC DBM1, M94110; AAC22364.1; ..
DR HSSP, P18400; ICYX.
DR InterPro; IPR001505; Copper_CuA.
DR InterPro; IPR008972; Cupredoxin.
DR InterPro; IPR002429; Cyt_c ox_2.
DR InterPro; IPR003088; Cyt_O1.
DR InterPro; IPR000345; CytC_heme_B5.
DR InterPro; IPR000437; Prok_LipoProt_S.
DR Pfam; PF00116; COX2_1.
DR Pfam; PF02790; COX2_TM; 1.
DR Pfam; PF00034; cytochrome_c; 1.
DR PRINTS; P00116; CYCOXIDASE11.
DR ProDom; PD000131; Copper_CuA; 1.
DR PROSITE; PS00013; PROKAR_LIPOPROTEIN; 1.
DR PROSITE; PS00078; COX2_1.
DR PROSITE; PS00190; CYTOCHROME_C; 1.
DR Oxidoreductase; Respiratory chain; Electron transport; Transmembrane;
DR Heme; Copper; Signal; Lipoprotein; Palmitate.
KM SIGNAL 1
FT CHAIN 23 342 CYTOCHROME C OXIDASE POLYPEPTIDE II.
FT DOMAIN 23 50 EXTRACELLULAR (POTENTIAL).
FT TRANSMEM 51 69 POTENTIAL.
FT DOMAIN 70 89 CYTOPLASMIC (POTENTIAL).
FT TRANSSEM 90 108 POTENTIAL.
FT DOMAIN 109 342 EXTRACELLULAR (POTENTIAL).
FT Lipid 23 23 N-palmitoyl cysteine (potential).
FT Lipid 23 23 S-diacylglycerol cysteine (potential).
FT METAL 175 175 COPPER A (PROBABLE).
FT METAL 210 210 COPPER A (PROBABLE).
FT METAL 214 214 COPPER A (PROBABLE).
FT METAL 218 218 COPPER A (PROBABLE).
FT BINDING 264 264 HEME (COVALENT) (PROBABLE).
FT BINDING 267 267 HEME (COVALENT) (PROBABLE).
FT METAL 268 268 IRON (HEME AXIAL LIGAND) (PROBABLE).
FT METAL 317 317 IRON (HEME AXIAL LIGAND) (PROBABLE).
SQ SEQUENCE 342 AA; 38141 MW; F777C7D1FE44A429 CRC64;

Query Match 31.8%; Score 48; DB 1; Length 342;
Best Local Similarity 23.3%; Pred. No. 9.7;
Matches 7; Conservative 10; Mismatches 13; Indels 0; Gaps 0;

Qy 1 DEWYSLAVDAQHAKRVASGLRFFELNA 30
Db 234 DAWVEGMSAEVPEPTETLANQGRQVFEENS 263

RESULT 7
WSB1_MOUSE STANDARD; PRT; 421 AA.
AC OS04927;
DT 28-FEB-2003 (Rel. 41, Created)
DT 28-FEB-2003 (Rel. 41, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE WP Repeat and SOCS box containing protein 1 (WSB-1).
GN WSB1.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
CX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=NIH Swiss;
RX MEDLINE=98081836; PubMed=9419338;
RA Hilton D.J., Richardson R.T., Alexander W.S., Viney E.M.,
RA Willson T.A., Spriggs N.S., Stair R., Nicholson S.E., Metcalf D.,
RA Nicola N.A.;
RT "Twenty proteins containing a C-terminal SOCS box form five structural
classes.";
```

```

RL Proc. Natl. Acad. Sci. U.S.A. 95:114-119 (1998).
CC -1- SIMILARITY: Contains 5 WD repeats.
CC -1- SIMILARITY: Contains 1 SOCS box domain.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL: AF03186; AAB96647.1; -.
CC MGD; MGI:1926139; Mabl.
CC InterPro: IPR001496; SOCS_C.
CC InterPro: IPR001680; WD40.
CC Pfam; PF00400; WD40; 6.
CC PRINTS; PR00320; GPROTEINRPT.
CC PRODOM; PD000018; WD40; 2.
CC SMART; SM00253; SOCS; 1.
CC SMART; SM00320; WD40; 6.
CC PROSITE; PS00678; WD_REPEATS_1; 2.
CC PROSITE; PS00682; WD_REPEATS_2; 6.
CC PROSITE; PS0294; WD_REPEATS_REGION; 1.
CC PROSITE; PS0225; SOCS; 1.
CC Repeat; WD repeat.
CC KW REPEAT 124 168 WD 1.
CC FT REPEAT 168 205 WD 2.
CC FT REPEAT 212 251 WD 3.
CC FT REPEAT 254 293 WD 4.
CC FT REPEAT 309 346 WD 5.
CC FT DOMAIN 372 421 WD 5.
CC SQ SEQUENCE 421 AA; 47065 MW; 4003D1FPE7A9A2EP CRC64;

Query Match 31.8%; Score 48; DB 1; Length 421;
Best Local Similarity 35.5%; Pred. No. 12;
Matches 11; Conservative 6; Mismatches 12; Indels 2; Gaps 1;

OY 1 DEWVR-SLAVDAQAKRVASGLRFRRLN 29
DB 311 DRWVAVSFSDHGLHVASLADKRVFRIRID 341

RESULT 8
SACL_CHLRE STANDARD; PRT; 585 AA.
ID SACL_CHLRE STANDARD; PRT; 585 AA.
AC Q39593;
DT 01-NOV-1997 (Rel. 35, Created)
DT 01-NOV-1997 (Rel. 35, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Putative sulfur deprivation response regulator.
GN SACL.
OS Chlamydomonas reinhardtii.
OC Eukaryota; Viridiplantae; Chlorophyta; Chlorophyceae; Volvocales;
OC Chlamydomonadales; Chlamydomonadales; Chlamydomonadales.
OC NCBI_TaxID=3055;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=137c / CC-125;
RX MEDLINE=96208501; PubMed=8641280;
RA Davies J.P., Yildiz F.H., Grossman A.;
RT "Sacl", a putative regulator that is critical for survival of
RT Chlamydomonas reinhardtii during sulfur deprivation."
RL EMBL J 15:2150-2159(1996).
CC -1- FUNCTION: Not known. Mutations in SACL produces cells that cannot
CC synthesize arylsulfatase and cannot take up sulfate as rapidly as
CC wild-type cells. SACL is necessary for cells to survive sulfur
CC deprivation.
CC -1- SUBCELLULAR LOCATION: Integral membrane protein (potential).
CC -1- SIMILARITY: BELONGS TO THE SACL3A FAMILY OF TRANSPORTERS.
CC NACD SUBFAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration

```

```

CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL: U47541; AAB08008.1; -.
CC PIR; S69216; S69216.
CC InterPro: IPR001898; Na/sul_symport.
CC InterPro: IPR006037; TrkA_C.
CC Pfam; PF02080; TrkA-C_1.
CC PROSITE; PS01271; NA_SULFATE; 1.
CC KW Transmembrane; Transport.
CC FT TRANSMEM 5 25 POTENTIAL.
CC FT TRANSMEM 30 50 POTENTIAL.
CC FT TRANSMEM 83 103 POTENTIAL.
CC FT TRANSMEM 117 137 POTENTIAL.
CC FT TRANSMEM 162 182 POTENTIAL.
CC FT TRANSMEM 389 409 POTENTIAL.
CC FT TRANSMEM 411 431 POTENTIAL.
CC FT TRANSMEM 442 462 POTENTIAL.
CC FT TRANSMEM 482 502 POTENTIAL.
CC FT TRANSMEM 561 581 POTENTIAL.
CC SQ SEQUENCE 585 AA; 62329 MW; 0603B05D86BD81C3 CRC64;

Query Match 31.8%; Score 48; DB 1; Length 585;
Best Local Similarity 42.9%; Pred. No. 17;
Matches 12; Conservative 5; Mismatches 9; Indels 2; Gaps 1;

OY 2 EWRSLAVDAQ-HAKRVASGLRFR 27
DB 199 EYVSVAVDARFAHIGRTIESAGLRHLR 226

RESULT 9
ABS_DROME STANDARD; PRT; 619 AA.
ID ABS_DROME STANDARD; PRT; 619 AA.
AC Q9V3C0; Q9V6D0;
DT 16-OCT-2001 (Rel. 40, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DT 10-OCT-2003 (Rel. 42, Last annotation update)
DE DEAD-box protein abstract.
GN ABS OR CG14637.
OS Drosophila melanogaster (Fruit fly).
OC Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
OC Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
OC Ephydroidea; Drosophilidae; Drosophila.
OC NCBI_TaxID=7227;
RN [1]
RP SEQUENCE FROM N.A.
RC MEDLINE=20171041; PubMed=10704843;
RX Schmucker D., Vorduegen G., Yeghiayan P., Fan H.Q., Jaekle H.,
RA Gaul U.;
RT "The Drosophila gene abstract, required for visual system development,
RT encodes a putative RNA helicase of the DEAD box protein family."
RL Mech. Dev. 91:189-196(2000).
RN [2]
RP SEQUENCE FROM N.A. AND VARIANTS 1AB AND 33B.
RC MEDLINE=20076860; PubMed=10607561;
RX Iltis U., Leptin M.;
RT "Developmental and cell biological functions of the Drosophila DEAD-
RT box protein abstract."
RL Curr. Biol. 9:1373-1381(1999).
RN [3]
RP SEQUENCE FROM N.A.
RC STRAIN=Berkeley;
RX MEDLINE=20196006; PubMed=10731132;
RA Adams M.D., Celiker S.E., Holt R.A., Evans C.A., Gocayne J.D.,
RA Amanatides P.G., Scherer S.E., Li P.W., Hoskins R.A., Galle R.F.,
RA George R.A., Lewis S.E., Richards S., Ashburner M., Henderson S.N.,
RA Sutton G.G., Wortman J.R., Vandal M.D., Zhang Q., Chen L.X.,
RA Brandon R.C., Rogers Y.-H.C., Blazey R.C., Champe M., Pfeiffer B.D.,

```



```

CC similarity).
CC -1- CATALYTIC ACTIVITY: S-adenosyl-L-methionine + 2-polyprenyl-6-
CC methoxy-1,4-benzoquinol = S-adenosyl-L-homocysteine + 2-
CC polyprenyl-3-methyl-6-methoxy-1,4-benzoquinol.
CC -1- CATALYTIC ACTIVITY: S-adenosyl-L-methionine + demethylmenaquinol =
CC S-adenosyl-L-homocysteine + menaquinol.
CC -1- PATHWAY: Menaquinone biosynthesis; last step.
CC -1- PATHWAY: Ubiquinone biosynthesis.
CC -1- SIMILARITY: Belongs to the ubiE family.
-----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
-----
CC EMBL; AE009003; AAL41342.1; ALT_INIT.
CC EMBL; AE007970; AAK6136.1; -.
CC PIR; G97397; G97397.
CC HAMAP; MF_01813; -.
CC InterPro; IPR001601; Methyltransf.
CC InterPro; IPR000051; SAM bind.
CC InterPro; IPR004034; Ubi/men Mentransf.
CC InterPro; IPR004033; UbiE/COG5 Metrf.
CC Pfam; PF01209; UbiE_methyltransf. 1.
CC PROSITE; PS01183; UBI_E_1.
CC PROSITE; PS01184; UBI_E_2; FALSE_NEG.
CC Menaquinone biosynthesis; Ubiquinone biosynthesis; Transferase;
CC Methyltransferase; Complete proteome.
CC SEQUENCE 259 AA; 28862 MW; 4F017ED69E2C7FFB CRC64;

Query Match 31.1%; Score 47; DB 1; Length 259;
Best Local Similarity 45.8%; Pred. No. 10;
Matches 11; Conservative 3; Mismatches 10; Indels 0; Gaps 0;

QY 7 LAVDAQAAKRVASGLRFFRLNA 30
DB 109 LAVGERAAKKLSNLTFFVEANA 132

RESULT 11
FAH2 STRAW STANDARD; PRT; 315 AA.
AC Q82M29;
DT 10-OCT-2003 (Rel. 42; Created)
DT 10-OCT-2003 (Rel. 42; Last sequence update)
DT 10-OCT-2003 (Rel. 42; Last annotation update)
DE 3-oxoacyl-(acyl-carrier-protein) synthase III protein 2 (EC 2.3.1.41)
DE (Beta-ketolacyl-ACP synthase III 2) (KAS III 2).
GN FAH2 OR SAV1831.
OS Streptomyces avermitilis.
OC Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;
OC Streptomyces; Streptomycetaceae; Streptomyces.
OX NCBI_TaxID=33903;

SEQUENCE FROM N.A.
RP STRAIN=MA-4680 / ATCC 31267 / NCIMB 12804 / NRRL 8165;
RC MEDLINE=21477403; PubMed=11572948;
RA Omura S., Ikeda H., Ishikawa U., Hanamoto A., Takahashi C.,
RA Shinose M., Takahashi Y., Horikawa H., Nakazawa H., Osone T.,
RA Kikuchi H., Shiba T., Sakaki Y., Hatori M.;
RT "Genome sequence of an industrial microorganism Streptomyces
RT avermitilis: deducing the ability of producing secondary
RT metabolites."
RT Proc. Natl. Acad. Sci. U.S.A. 98:12215-12220(2001).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=MA-4680 / ATCC 31267 / NCIMB 12804 / NRRL 8165;
RC MEDLINE=22608306; PubMed=12692562;
RA Ikeda H., Ishikawa U., Hanamoto A., Shinose M., Kikuchi H., Shiba T.,

```

```

RA Sakaki Y., Hatori M., Omura S.;
RT "Complete genome sequence and comparative analysis of the industrial
RT microorganism Streptomyces avermitilis.";
RL Nat. Biotechnol. 21:526-531(2003).
CC -1- FUNCTION: Catalyzes the condensation reaction of fatty acid
CC synthesis by the addition to an acyl acceptor of two carbons from
CC malonyl-ACP. Catalyzes the first condensation reaction which
CC initiates fatty acid synthesis and may therefore play a role in
CC governing the total rate of fatty acid production. Possesses both
CC acetoacetyl-ACP synthase and acetyl transacylase activities. Its
CC substrate specificity determines the biosynthesis of branched-
CC chain and/or straight-chain of fatty acids (By similarity).
CC -1- CATALYTIC ACTIVITY: Acyl-(acyl-carrier protein) + malonyl-(acyl-
CC carrier protein) = 3-oxoacyl-(acyl-carrier protein) + CO(2) +
CC [acyl-carrier protein].
CC -1- PATHWAY: Fatty acid biosynthesis.
CC -1- SUBUNIT: Homodimer (By similarity).
CC -1- DOMAIN: The last Arg residue of the ACP-binding site is essential
CC for the weak association between ACP/acpP and fabH (By
CC similarity).
CC -1- SIMILARITY: Belongs to the fabH family.
-----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
-----
CC EMBL; AP005028; BAC69542.1; -.
CC HAMAP; MF_01815; -.
CC InterPro; IPR004655; FabH.
CC TRIGRAMS; TIGR00747; fabH; 1.
CC Fatty acid biosynthesis; Transferase; Acyltransferase;
CC Multifunctional enzyme; Complete proteome.
CC FT ACT_SITE 113 113 BY SIMILARITY.
CC FT ACT_SITE 241 241 BY SIMILARITY.
CC FT ACT_SITE 271 271 BY SIMILARITY.
CC FT SITE 242 246 ACP-BINDING (BY SIMILARITY).
CC SEQUENCE 315 AA; 33371 MW; E490D12CBA86583C CRC64;

Query Match 31.1%; Score 47; DB 1; Length 315;
Best Local Similarity 33.3%; Pred. No. 13;
Matches 13; Conservative 5; Mismatches 5; Indels 16; Gaps 2;

QY 1 DEWVRS-LAVDAQ-----AAKRVASGL 23
DB 31 DEWIRSVGIRTRHFGAPDEPVDELGAHAAKALASAGL 69

RESULT 12
DCUP XANAC STANDARD; PRT; 354 AA.
AC O8P185;
DT 28-FEB-2003 (Rel. 41; Created)
DT 28-FEB-2003 (Rel. 41; Last sequence update)
DT 28-FEB-2003 (Rel. 41; Last annotation update)
DE Uroporphyrinogen decarboxylase (EC 4.1.1.37) (URO-D) (UPD).
GN HEME OR XAC013.
OS Xanthomonas axonopodis (pv. citri).
OC Bacteria; Proteobacteria; Gammaproteobacteria; Xanthomonadales;
OC Xanthomonadaceae; Xanthomonas.
OX NCBI_TaxID=92829;

SEQUENCE FROM N.A.
RP STRAIN=306 / ATCC 13902 / XV 101;
RC MEDLINE=22022145; PubMed=12024217;
RA da Silva A.C.R., Perro J.A., Reinach P.C., Farah C.S., Furlan L.R.,
RA Quaglini R.B., Monteiro-Vitorello C.B., Van Sluys M.A., Almeida N.F.,
RA Alves L.M.C., do Amaral A.M., Bertolini M.C., Camargo L.B.A.,
RA Camarotte G., Camarvan F., Cardozo C., Chambergo F., Clapina L.F.,

```


OC Methanocaldococcaceae; Methanocaldococcus.
 OX NCBI_TaxID=2190;
 RN [1]
 RC SEQUENCE FROM N.A.
 RA STRAIN=JAL-1 / DSM 2661 / ATCC 43067;
 RX MEDLINE=96337999; PubMed=668087;
 RA But C.G., White O., Olsen G.J., Zhou L., Fleischmann R.D.,
 RA Sutton G.G., Blake J.A., Fitzgerald J.M., Clayton R.A., Gocayne J.D.,
 RA Kerlavage A.R., Dougherty B.A., Tomb J.-F., Adams M.D., Reisch C.I.,
 RA Overbeek R., Kirkness E.F., Weinstock K.G., Merrick J.M., Glodek A.,
 RA Scott J.L., Geoghegan N.S.M., Weidman J.F., Ruhnann J.L., Nguyen D.,
 RA Uettermann T.R., Kelley J.M., Peterson J.D., Sadow P.W., Hanna M.C.,
 RA Cotton M.D., Roberts K.M., Hurst M.A., Kaine B.F., Borodovsky M.,
 RA Klenk H.-P., Fraser C.M., Smith H.O., Woese C.R., Venter J.C.;
 RA "Complete genome sequence of the methanogenic archaeon, Methanococcus
 jannaschii";
 RT Science 273:1058-1073(1996).
 RL Science 273:1058-1073(1996).
 CC -1- FUNCTION: Involved in protein export. Interacts with secA and secE
 CC to allow the translocation of proteins across the plasma membrane,
 CC by forming part of a channel (By similarity).
 CC -1- SUBCELLULAR LOCATION: Integral membrane protein.
 CC -1- SIMILARITY: Belongs to the secY/SEC1-alpha family.
 CC
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC
 CC -----
 CC EMBL: U67498; AAB98469.1; -;
 DR TIGR: MJ0478; -;
 DR InterPro: IPR002208; SecY.
 DR Pfam: PF00344; secy: 1.
 DR PRINTS: PR00303; SECYRLCASE.
 DR TIGRPFAMS: TIGR00967; 3A05018007; 1.
 DR PROSITE: PS00785; SECY_2; 1.
 DR PROSITE: PS00756; SECY_2; 1.
 DR KX protein transport; Transmembrane; Translocation; Complete proteome.
 FT TRANSMEM 29 49 POTENTIAL.
 FT TRANSMEM 70 90 POTENTIAL.
 FT TRANSMEM 125 145 POTENTIAL.
 FT TRANSMEM 166 186 POTENTIAL.
 FT TRANSMEM 209 229 POTENTIAL.
 FT TRANSMEM 257 277 POTENTIAL.
 FT TRANSMEM 314 334 POTENTIAL.
 FT TRANSMEM 384 404 POTENTIAL.
 SQ SEQUENCE 436 AA; 47444 MW; 3B67747A7B5BDC0A CRC64;

Query Match 30.5%; Score 46; DB 1; Length 436;
 Best Local Similarity 38.1%; Pred. No. 25;
 Matches 8; Conservative 5; Mismatches 8; Indels 0; Gaps 0;

QY 3 WYSLAVDAQHAKRVASEGL 23
 Db 334 WETITGLDPKSMKAKRIGSLGM 354

RESULT 15
 DONS DROME STANDARD; PRT; 568 AA.
 AC Q9VNA8;
 DT 16-OCT-2001 (Rel. 40, Created)
 DT 16-OCT-2001 (Rel. 40, Last sequence update)
 DT 10-OCT-2003 (Rel. 42, Last annotation update)
 DB Downstream of son gene protein homolog.
 GN CG2669.
 OS Drosophila melanogaster (Fruit fly).
 OC Insecta; Pterygota;
 OC Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
 OC Ephydroidea; Drosophilidae; Drosophila.
 OX NCBI_TaxID=7227;

RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=Berkley;
 RX MEDLINE=20196006; PubMed=10731132;
 RA Adams M.D., Celinker S.E., Holt R.A., Evans C.A., Gocayne J.D.,
 RA Amaralides P.G., Scherer S.E., Li P.W., Hoekins R.A., Galle R.F.,
 RA George R.A., Lewis S.E., Richards S., Ashburner M., Henderson S.N.,
 RA Sutton G.G., Mortan J.R., Yandell M.D., Zhang Q., Chen L.X.,
 RA Brandon R.C., Rogers Y.-H.C., Blazer R.G., Chapple M., Pfeiffer B.D.,
 RA Wan K.H., Doyle C., Baxter E.G., Helt G., Nelson C.R., Miklos G.L.G.,
 RA Abil J.F., Agbanyani A., An H.-J., Andrews-Pfannkoch C., Baldwin D.,
 RA Bailew R.M., Basu A., Baxendale J., Bayraktoglu L., Beasley E.M.,
 RA Beeson K.Y., Benos P.V., Berman B.P., Bhandari D., Bolshakov S.,
 RA Borokova D., Botchan M.R., Boulter J., Brockstein P., Brotter P.,
 RA Burtis K.C., Busan D.A., Butler C., Cadieu E., Center A., Chandra I.,
 RA Cherry J.M., Cawley S., Dahlke C., Davaport L.B., Davies P.,
 RA de Pablo B., Delcher A., Deng Z., Mays A.D., Dew I., Dietz S.M.,
 RA Dodson K., Doup L.E., Downes M., Dugan-Rocha S., Dunkov B.C., Dunn P.,
 RA Durbin K.J., Evangelista C.C., Ferraz C., Ferreira S., Fleischmann W.,
 RA Foster C., Gabrielian A.B., Garg N.S., Gelbart W.M., Glasser K.,
 RA Glodek A., Gong F., Gorrell J.H., Gu Z., Guan P., Harris M.,
 RA Harris N.L., Harvey D.A., Heiman T.J., Hernandez J.R., Houck J.,
 RA Hostin D., Houston K.A., Howland T.J., Wei M.-H., Ibegwan C.,
 RA Ujalji M., Kalush F., Karpen G.H., Ke Z., Kennison J.A., Ketchum K.A.,
 RA Kimmel B.E., Kodira C.D., Kraft C., Kravitz S., Kuip D., Lai Z.,
 RA Lasko P., Lei Y., Levitky A.A., Li J.H., Li Z., Liang Y., Lin X.,
 RA Liu X., Mattei B., McIntosh T.C., McLeod M.P., McPherson D.,
 RA Merkulov G., Mishina N.V., Mobarry C., Morris J., Moshrefi A.,
 RA Mount S.M., Moy M., Murphy B., Murphy L., Muzny D.N., Nelson D.L.,
 RA Nelson D.R., Nelson K.A., Nixon K., Nusskern D.R., Paley J.M.,
 RA Palazolo M., Pittman G.S., Pan S., Pollard J., Puti V., Reese M.G.,
 RA Reinert K., Remington K., Saunders R.D.C., Scheeler F.P., Shen H.,
 RA Shue B.C., Siden-Kiamos I., Simpson M., Strong R., Sun E.,
 RA Slater E., Spradling A.C., Stapleton M., Strong R., Sun E.,
 RA Svirskas R., Tector C., Turner R., Venter E., Wang A.H., Wang X.,
 RA Wang Z.-Y., Wasserman D.A., Weinstock G.M., Weissbach J.,
 RA Williams S.M., Woodage T., Worley K.C., Wu D., Yang S., Yao Q.A.,
 RA Ye J., Yeh R.-F., Zaveri J.S., Zhan M., Zhang G., Zhao Q., Zheng L.,
 RA Zheng X.H., Zhong F.N., Zhong W., Zhou X., Zhu S., Zhu X., Smith H.O.,
 RA Gibbs R.A., Myers E.W., Rubin G.M., Venter J.C.;
 RA "The genome sequence of Drosophila melanogaster";
 RT Science 287:2185-2195(2000).
 RL Science 287:2185-2195(2000).
 CC -1- SIMILARITY: Belongs to the DONSON family.
 CC
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC
 CC -----
 CC EMBL: AB003603; AAF52038.1; -;
 DR FLYBASE: FBgn0037316; CG2669.
 KW Hypothetical protein.
 SQ SEQUENCE 568 AA; 63437 MW; A7E29D4C0AFD1A80 CRC64;

Query Match 30.5%; Score 46; DB 1; Length 568;
 Best Local Similarity 34.5%; Pred. No. 33;
 Matches 10; Conservative 6; Mismatches 5; Indels 8; Gaps 1;

QY 1 DEWYSLAVD-----AQHAKRVASE 21
 Db 355 EDWLESIGVDERTIRIQTSSHARKQAAE 363

Search completed: September 30, 2004, 11:10:33
 Job time : 13.1429 secs

This Page Blank (uspto)

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: September 30, 2004, 10:58:43 / Search time 19.1429 Seconds
(without alignments)
150.748 Million cell updates/sec

Title: US-09-988-851A-11
Perfect score: 151
Sequence: 1 DEWRSILAVDAQHAKRVASGLRFFRLNA 30

Scoring table: BLOSUM62
Gapop 10.0, Gapext 0.5

Searched: 283366 seqs, 96191526 residues
Total number of hits satisfying chosen parameters: 283366

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database: PIR 78:1
1: PIR1:1
2: PIR2:1
3: PIR3:1
4: PIR4:1

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	151	100.0	798	2	S09867
2	55	36.4	1998	2	T08822
3	53	35.1	1318	1	H1BPD7
4	52	34.4	429	2	A70949
5	51	33.8	421	2	T14773
6	50.5	33.4	757	2	B84790
7	50	33.1	347	2	I64213
8	49	32.5	407	2	G84783
9	49	32.5	1533	2	T00344
10	49	32.5	3498	2	T22330
11	48	31.8	342	2	C45335
12	48	31.8	585	2	S69216
13	48	31.8	1871	2	S27938
14	47.5	31.5	473	2	E72723
15	47	31.1	258	2	AH6315
16	47	31.1	259	2	G97397
17	47	31.1	295	2	S67934
18	47	31.1	454	2	T36209
19	47	31.1	612	2	T50226
20	47	31.1	1524	2	T30518
21	46.5	30.8	232	2	A83208
22	46	30.5	113	2	A95909
23	46	30.5	394	2	C82951
24	46	30.5	440	1	F64359
25	46	30.5	581	2	F82723
26	46	30.5	670	2	AH1664
27	46	30.5	1242	2	S78061
28	45.5	30.1	502	1	B40620
29	45.5	30.1	502	2	D86791

30	45	29.8	117	2	A72692	hypothetical prote
31	45	29.8	185	2	F75549	cytochrome c bioge
32	45	29.8	200	2	T02173	probable alcohol d
33	45	29.8	255	1	XVECG1	tRNA (guanine-N1-)
34	45	29.8	255	2	S31775	tRNA (guanine-N1-)
35	45	29.8	255	2	B85907	tRNA (guanine-N1-)
36	45	29.8	255	2	F91062	tRNA (guanine-N1-)
37	45	29.8	255	2	AF0833	2-dehydro-3-deoxyg
38	45	29.8	295	2	AG2858	hypothetical prote
39	45	29.8	301	2	D87508	hypothetical prote
40	45	29.8	334	2	E97635	hypothetical prote
41	45	29.8	361	2	B96016	hypothetical prote
42	45	29.8	370	2	AB3175	hypothetical prote
43	45	29.8	493	2	E69487	protein translocas
44	45	29.8	516	2	AB3438	amidophosphoribosy
45	45	29.8	536	2	B81319	N-acetylneuraminate

ALIGNMENTS

RESULT 1
S09867
hypothetical protein UH102 - human cytomegalovirus (strain AD169)
C/Species: human cytomegalovirus, human herpesvirus 5
A/Note: host Homo sapiens (man)
C/Date: 07-Sep-1990 #sequence_revision 07-Sep-1990 #text_change 08-Oct-1999
C/Accession: S09867
R/Chae, M.S.; Bankier, A.T.; Beck, S.; Bohni, R.; Brown, C.M.; Cerny, R.; Hornsnel, T.;
M.; Barrell, B.G.
Curr. Top. Microbiol. Immunol. 154, 125-169, 1990
A/Title: Analysis of the protein-coding content of the sequence of human cytomegalovirus
A/Reference number: S09749; MUID:90269039; PMID:2161319
A/Accession: S09867
A/Status: nucleic acid sequence not shown; translation not shown
A/Molecule type: DNA
A/Residues: 1-798 <CHE>
A/Cross-references: EMBL:X17403; NID:959591; PIDN:CAA3538.1; PID:e27303; PID:g1780881
A/Note: this sequence was submitted to the EMBL Data Library, December 1989
A/Note: this reading frame extends between two stop codons and does not begin with a str:

Query Match 100.0%; Score 151; DB 2; Length 798;

Best Local Similarity 100.0%; Pred. No. 5.7e-14; Mismatches 0; Indels 0; Gaps 0;

Query 1 DEWRSILAVDAQHAKRVASGLRFFRLNA 30
Db 769 DEWRSILAVDAQHAKRVASGLRFFRLNA 798

RESULT 2

T08822
nonstructural polypeptide - Rhopalosiphum padi virus
C/Species: Rhopalosiphum padi virus
C/Date: 20-Sep-1999 #sequence_revision 20-Sep-1999 #text_change 21-Jul-2000
C/Accession: T08822
R/Moon, J.S.; Domier, L.L.; McCoppin, N.K.; D'Arcy, C.J.; Jin, H.
Virology 243, 54-65, 1998
A/Title: Nucleotide sequence analysis shows that Rhopalosiphum padi virus is a member of
A/Reference number: Z16481; MUID:98201645; PMID:9527915
A/Accession: T08822
A/Status: translated from GB/EMBL/DDBT
A/Molecule type: genomic RNA
A/Residues: 1-1998 <MOO>
A/Cross-references: EMBL:AF022937; NID:92911298; PIDN:AA095509.1; PID:92911299
C/Keywords: polypeptide

Query Match 36.4%; Score 55; DB 2; Length 1998;
Best Local Similarity 50.0%; Pred. No. 16;
Matches 10; Conservative 5; Mismatches 5; Indels 0; Gaps 0;

Query 9 VDAQHAKRVASGLRFFRL 28
Db 769 DEWRSILAVDAQHAKRVASGLRFFRLNA 798

Db 501 LETOAGRMVSEGRIMRI 520

RESULT 3

Internal vitron protein D - phage T7

C/Species: phage T7

C/Date: 13-Jun-1993 #sequence_revision 13-Jun-1993 #text_change 23-Jul-1999

C/Accession: A04352; S42332

R/Date: 13-Jun-1993 #sequence_revision 13-Jun-1993 #text_change 23-Jul-1999

A/Reference number: A94615

A/Accession: A04352

A/Molecule type: DNA

A/Residues: 1-1318 <DUN>

R/Date: 13-Jun-1993 #sequence_revision 13-Jun-1993 #text_change 23-Jul-1999

A/Reference number: A94615

A/Accession: A04352

A/Molecule type: DNA

A/Residues: 1-1318 <DUN>

A/Cross-references: EMBL:V01146; NID:g431187; PIDN:CAA24434.1; PID:g15610

C/Map position: 76.61-86.51

C/Superfamily: phage T7 internal vitron protein D

Query Match

Best Local Similarity 35.1%; Score 53; DB 1; Length 1318;

Matches 12; Conservative 3; Mismatches 5; Indels 0; Gaps 0;

Db 5 RSLAVDAQHAKRVASEGLR 24

358 KALVVGASALNVAASEGLR 377

RESULT 4

hypotheical protein RV3179 - Mycobacterium tuberculosis (strain H37RV)

C/Species: Mycobacterium tuberculosis

C/Date: 17-Jul-1998 #sequence_revision 17-Jul-1998 #text_change 20-Jun-2000

C/Accession: A70949

R/Date: 17-Jul-1998 #sequence_revision 17-Jul-1998 #text_change 20-Jun-2000

A/Reference number: A70949

A/Accession: A70949

A/Molecule type: DNA

A/Residues: 1-429 <COL>

A/Cross-references: GB:AL021646; GB:AL123456; NID:g3242278; PIDN:CAA16644.1; PID:g282758

A/Experimental source: strain H37RV

C/Genetics:

A/Status: Preliminary; nucleic acid sequence not shown; translation not shown

A/Map position: 2

Query Match

Best Local Similarity 34.4%; Score 52; DB 2; Length 429;

Matches 10; Conservative 3; Mismatches 7; Indels 0; Gaps 0;

Db 3 WRSALVDAQHAKRVASEGLR 22

53 WRSALVDAQHAKRVASEGLR 72

RESULT 5

hypotheical protein DKFZp564B0482.1 - human

C/Species: Homo sapiens (man)

C/Date: 20-Sep-1999 #sequence_revision 20-Sep-1999 #text_change 20-Sep-1999

C/Accession: T14773

R/Date: 20-Sep-1999 #sequence_revision 20-Sep-1999 #text_change 20-Sep-1999

A/Reference number: T14773

A/Accession: T14773

A/Molecule type: mRNA

A/Residues: 1-421 <POU>

A/Cross-references: EMBL:AL110243

A/Experimental source: fetal brain; clone DKFZp564B0482

C/Genetics:

A/Note: DKFZp564B0482.1

Query Match

Best Local Similarity 33.8%; Score 51; DB 2; Length 421;

Matches 11; Conservative 7; Mismatches 11; Indels 2; Gaps 1;

Db 1 DEWRSLAV--DAQHAKRVASEGLRFPRLN 29

311 DEWRSLAV--DAQHAKRVASEGLRFPRLN 341

RESULT 6

hypotheical protein At2g37230 (imported) - Arabidopsis thaliana

C/Species: Arabidopsis thaliana (mouse-ear cress)

C/Date: 02-Feb-2001 #sequence_revision 02-Feb-2001 #text_change 02-Feb-2001

C/Accession: B84790

R/Date: 02-Feb-2001 #sequence_revision 02-Feb-2001 #text_change 02-Feb-2001

A/Reference number: B84790

A/Accession: B84790

A/Molecule type: DNA

A/Residues: 1-757 <STO>

A/Cross-references: GB:AE002093; NID:g4056478; PIDN:AA098044.1; GSPDB:GNO0139

A/Experimental source: strain G-37

C/Genetics:

A/Status: Preliminary

A/Map position: 2

Query Match

Best Local Similarity 42.3%; Score 50.5; DB 2; Length 757;

Matches 11; Conservative 4; Mismatches 10; Indels 1; Gaps 1;

Db 2 EWRSALVDAQHAKRVASEGLR 27

114 EWRSALVDAQHAKRVASEGLR 138

RESULT 7

tryptophan-tRNA ligase (BC 6.1.1.2) - Mycoplasma genitalium

C/Species: Mycoplasma genitalium

C/Date: 17-Nov-1995 #sequence_revision 17-Nov-1995 #text_change 03-Jun-2002

C/Accession: I64213

R/Date: 17-Nov-1995 #sequence_revision 17-Nov-1995 #text_change 03-Jun-2002

A/Reference number: I64213

A/Accession: I64213

A/Molecule type: DNA

A/Residues: 1-347 <TIGR>

A/Cross-references: GB:U39691; GB:U43967; NID:g1045794; PID:g1045806; TIGR:MG126

A/Experimental source: strain G-37

C/Genetics:

A:Genetic code: SGC3

C:Superfamily: tryptophan-tRNA ligase
C:Keywords: aminoacyl-tRNA synthetase, ligase; protein biosynthesis

Query Match 33.1%; Score 50; DB 2; Length 347;
Best Local Similarity 42.3%; Pred. No. 14;
Matches 11; Conservative 5; Mismatches 10; Indels 0; Gaps 0;

Qy 1 DEWVSLAVDAQHAKRVASEGLRPF 26
Db 309 DEWVKVLDGKQAKVADETLKPF 334

RESULT 8

G84783
probable pectinesterase [imported] - Arabidopsis thaliana

C:Species: Arabidopsis thaliana (mouse-ear cress)
C>Date: 02-Feb-2001 #sequence_revision 02-Feb-2001 #text_change 02-Feb-2001

C/Accession: G84783

R/Lin, X.; Kaul, S.; Rounsley, S.D.; Shea, T.P.; Benito, M.I.; Town, C.D.; Fujii, C.Y.;

M.; Koo, H.; Moffat, K.S.; Cronin, L.A.; Shen, M.; Vankken, S.E.; Umayam, L.; Tallon, L.

euss, D.; Nieman, W.C.; White, O.; Eisen, J.A.; Salzberg, S.L.; Fraser, C.M.; Venter, J.

Nature 402, 761-768, 1999

A>Title: Sequence and analysis of chromosome 2 of the plant Arabidopsis thaliana.

A/Reference number: A84420; MUID:20083487; PMID:10617197

A/Accession: G84783

A>Status: preliminary

A/Molecule type: DNA

A/Residues: 1-407 <SNO>

A/Cross-references: GB:A8002093; NID:g4415916; PIDN:AAD20147.1; GSPDB:GN00139

C/Genetics:

A/Gene: At2g36710

A/Map position: 2

Query Match 32.5%; Score 49; DB 2; Length 407;
Best Local Similarity 26.9%; Pred. No. 24;
Matches 7; Conservative 9; Mismatches 10; Indels 0; Gaps 0;

Qy 1 DEWVSLAVDAQHAKRVASEGLRPF 26
Db 382 DEWLRHTNIVSEHTSKDIGDLGHY 407

RESULT 9

T00344
hypothetical protein KIAA0586 - human

C/Species: Homo sapiens (man)

C>Date: 01-Feb-1999 #sequence_revision 01-Feb-1999 #text_change 21-Jul-2000

C/Accession: T00344

R/Nagase, T.; Ishikawa, K.; Miyajima, N.; Tanaka, A.; Kotani, H.; Nomura, N.; Ohara, O.

DNA Res. 5, 31-39, 1998

A>Title: Prediction of the coding sequences of unidentified human genes. IX. The complete

A/Reference number: Z14086; MUID:98290545; PMID:9628581

A/Accession: T00344

A>Status: preliminary; translated from GB/EMBL/DBJ

A/Molecule type: mRNA

A/Residues: 1-1533 <NMG>

A/Cross-references: EMBL:AB011158; NID:g3043695; PIDN:BA25512.1; PID:g1043696

A/Experimental source: brain

C/Genetics:

A/Note: KIAA0586

Query Match 32.5%; Score 49; DB 2; Length 1533;
Best Local Similarity 31.0%; Pred. No. 93;
Matches 9; Conservative 7; Mismatches 13; Indels 0; Gaps 0;

Qy 1 DEWVSLAVDAQHAKRVASEGLRPFRLN 29
Db 526 DEWIKTISABIDELSRDYEQKRFPOKN 554

RESULT 10
T22330

hypothetical protein F47A4.2 - Caenorhabditis elegans

C/Species: Caenorhabditis elegans

C>Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 18-Feb-2000

C/Accession: T22330

R/Mortimore, B.

submitted to the EMBL Data Library, June 1995

A/Reference number: Z15549

A/Accession: T22330

A>Status: preliminary; translated from GB/EMBL/DBJ

A/Molecule type: DNA

A/Residues: 13438 <MTL>

A/Cross-references: EMBL:Z49888; PIDN:CAA90064.1; GSPDB:GN00028; CESP:F47A4.2

A/Experimental source: clone F47A4

C/Genetics:

A/Gene: CESP:F47A4.2

A/Map position: X

A/Introns: 36/3; 72/3; 147/2; 228/3; 261/2; 320/2; 847/3; 881/1; 1317/2; 1500/3; 1653/2;

Query Match 32.5%; Score 49; DB 2; Length 3498;
Best Local Similarity 40.9%; Pred. No. 2,2e+02;
Matches 9; Conservative 6; Mismatches 7; Indels 0; Gaps 0;

Qy 3 WVRSLAVDAQHAKRVASEGLR 24
Db 1336 WGRVAFVFRHAKALANAGLO 1357

RESULT 11

C45335
cytochrome-c oxidase (EC 1.9.3.1) chain II - Bacillus firmus

C/Species: Bacillus firmus

C>Date: 30-Apr-1993 #sequence_revision 18-Nov-1994 #text_change 18-Aug-2003

C/Accession: C45335; I39821

R/Quirk, P.G.; Hicks, D.B.; Kruwich, T.A.

J. Biol. Chem. 268, 678-685, 1993

A>Title: Cloning of the cta operon from alkaliphilic Bacillus firmus OF4 and characteriz

A/Reference number: A45335; MUID:93107080; PMID:7678007

A/Accession: C45335

A/Molecule type: DNA; protein

A/Residues: 1-342 <QUT>

A/Cross-references: GB:M94110; NID:g142782; PIDN:AAA2364.1; PID:g142786

A/Experimental source: strain OF4

C/Genetics:

A/Gene: ctaC

C/Superfamily: cytochrome-c oxidase chain II, mitochondrial type; cytochrome-c oxidase c

C/Keywords: copper; electron transfer; heme; membrane-associated complex; oxidoreductase

C/Species: Chlamydomonas reinhardtii

F/15-228/Domain: cytochrome-c oxidase chain II homology <CO2>

F/115-210/214,221/Binding site: copper 1 (His, Cys, Met) #status predicted

F/210,212,214,218/Binding site: copper 2 (Cys, Glu, Cys, His) #status predicted

F/212/Binding site: magnesium (Glu) (shared with chain I) #status predicted

Query Match 31.8%; Score 48; DB 2; Length 342;
Best Local Similarity 23.3%; Pred. No. 28;
Matches 7; Conservative 10; Mismatches 13; Indels 0; Gaps 0;

Qy 1 DEWVSLAVDAQHAKRVASEGLRPFRLNA 30
Db 234 DAWEGMSAEVEPTETLANQGRQVEENS 263

RESULT 12

S69216

sulfur deprivation response regulator sac1 - Chlamydomonas reinhardtii

C/Species: Chlamydomonas reinhardtii

C>Date: 28-Oct-1996 #sequence_revision 13-Mar-1997 #text_change 21-Jul-2000

C/Accession: S69216

R/Davies, J.P.; Yildiz, F.H.; Grossman, A.

EMBO J. 15, 2150-2159, 1996

A>Title: Sac1, a putative regulator that is critical for survival of Chlamydomonas reinh.

A/Reference number: S69216; MUID:96208501; PMID:8641280

A/Accession: S69216

A>Status: preliminary; nucleic acid sequence not shown

A/Molecule type: mRNA

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: September 30, 2004, 10:58:43 / Search time 77.7143 Seconds
(without alignments)
109.072 Million cell updates/sec

Title: US-09-988-851A-11
Perfect score: 151
Sequence: 1 DEWRSILAVDAQHAKRVASEGLRFFRLNA 30

Scoring table: BLOSUM62
Gapop 10.0, Gapext 0.5

Searched: 1586107 seqs, 282547505 residues

Total number of hits satisfying chosen parameters: 1586107

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database: A_Geneseq_29Jan04:*
1: Geneseqp1980s:*
2: Geneseqp1990s:*
3: Geneseqp2000s:*
4: Geneseqp2001s:*
5: Geneseqp2002s:*
6: Geneseqp2003as:*
7: Geneseqp2003bs:*
8: Geneseqp2004s:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	151	100.0	30	2	AAW53267 Herpes si
2	97	64.2	26	2	AAW53268 Herpes si
3	52.5	34.8	873	4	AAU29536 Novel hum
4	51	33.8	221	6	AAU60589 Human NOR
5	51	33.8	221	6	ABU11555 Human MDR
6	51	33.8	312	2	AAW93276 Murine WD
7	51	33.8	345	2	AAW18879 Leucine a
8	51	33.8	389	2	AAW39367 SBHWSB1 c
9	51	33.8	421	2	AAW39368 SBHWSB1 c
10	51	33.8	421	3	AAH12445 Human MSB
11	51	33.8	421	3	AAW53885 A suppres
12	51	33.8	421	4	AAW93410 Human pol
13	51	33.8	421	4	AAW39188 Human pol
14	51	33.8	611	5	AAU82968 Human hom
15	51	33.8	874	4	AAU40974 Human pol
16	50	33.1	233	6	ADA33335 Acinetoba
17	50	33.1	347	6	ABU35532 Protein e
18	49.5	32.8	3745	6	AAH37601 Micromono
19	49.5	32.8	4471	6	AAH36999 Micromono
20	49	31.5	407	5	ABH92012 Herbicida
21	48	31.8	314	4	AAH60867 Mevalonat
22	48	31.8	314	6	ABU97197 Enzyme po
23	48	31.8	421	2	AAW62618 Mus muscu
24	48	31.8	421	3	AAH12446 Mouse MSB
25	48	31.8	1871	2	AAU1078 Sugarcane

26	48	31.8	1871	2	AAH15671 Sugarcane
27	48	31.8	1871	2	AAW57165 Amino aci
28	48	31.8	1871	6	ABU08544 Sugarcane
29	47.5	31.5	619	4	ABH69630 Drosophill
30	47	31.1	1716	6	ABH68549 Photorhab
31	46	30.5	255	3	ABH07035 HIV-2 GH
32	46	30.5	267	6	ABH67151 Photorhab
33	46	30.5	356	4	ABH63969 Drosophill
34	46	30.5	568	4	ABH59306 Drosophill
35	46	30.5	564	6	ABH57682 Saccharop
36	46	30.5	568	2	AAW39301 Spn2 a po
37	46	30.5	568	4	AAH70969 S. spins
38	45.5	30.1	251	5	ABH45400 Human Bly
39	45.5	30.1	502	5	ABH54671 Lactococc
40	45	29.8	113	6	ABU31356 Protein e
41	45	29.8	155	3	AAW91069 Streptomy
42	45	29.8	200	5	ABH92146 Herbicida
43	45	29.8	255	4	AAU34668 E. coli c
44	45	29.8	255	4	AAU38492 Salmonell
45	45	29.8	255	6	ABU47891 Protein e

ALIGNMENTS

RESULT 1	AAW53267	standard; peptide; 30 AA.
XX	AAW53267	
AC	AAW53267	
XX	01-JUL-1998	(first entry)
XX	Herpes simplex virus type 1	antiviral agent peptide 1.
XX	Herpes simplex virus type 1	HSV-1; antiviral; inhibition; infection;
KM	replication; UL8; POL.	
XX	Synthetic.	
OS	Herpes simplex virus unknown type.	
XX	WO9804707-A1.	
PD	05-FEB-1998.	
XX	28-JUL-1997;	97NO-GB002025.
XX	26-JUL-1996;	96GB-00015730.
XX	(MED1-) MEDICAL RES COUNCIL.	
PI	Marsden HS, Stow ND, McLean GW;	
XX	WPI; 1998-100695/12.	
XX	Antiviral agent capable of inhibiting herpes virus replication - disrupts	
PT	association between herpes simplex virus type 1 UL8 and POL.	
XX	Claim 5; Page 57; 83pp; English.	
XX	The present sequence represents an antiviral agent peptide capable of	
CC	combating herpes virus replication. The antiviral agent is capable of	
CC	disrupting the association between UL8 and POL (UL30), where UL8 and POL	
CC	are respectively defined as UL8 and POL of herpes simplex virus type 1	
CC	(HSV-1) together with homologues in other herpes viruses. The present	
CC	invention also describes an assay to determine the ability of a test	
CC	substance to interfere with the association of UL8 and POL. The assay	
CC	comprises: (a) exposing a 1st viral component to a test substance	
CC	followed by a 2nd viral component, or exposing a 1st viral component to a	
CC	2nd viral component followed by a test substance; (b) washing to remove	
CC	any 2nd viral component and/or test substance not associated with the 1st	
CC	viral component; and (c) detecting the presence, and optionally	
CC	determining the amount, of 2nd viral component associated with the 1st	

CC viral component. The antiviral agent can be used to combat herpes virus
CC replication or infection
XX
SQ Sequence 30 AA;

Query Match 100.0%; Score 151; DB 2; Length 30;
Best Local Similarity 100.0%; Pred. No. 3.6e-17;
Matches 30; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 DEWVRSLLAVDAQHAKRVASGGLRFRRLNA 30
1 DEWVRSLLAVDAQHAKRVASGGLRFRRLNA 30

RESULT 2
AAW53268
ID AAW53268 standard; peptide; 26 AA.

XX AAW53268;

DT 01-JUL-1998 (first entry)

DE Herpes simplex virus type 1 antiviral agent peptide 2.

XX Herpes simplex virus type 1, HSV-1; antiviral; inhibition; infection;

KW replication; UL8; POL.

XX Synthetic.

OS Herpes simplex virus unknown type.

PN W09804707-A1.

XX 05-FEB-1998.

PF 28-JUL-1997; 97WO-GB002025.

XX 26-JUL-1996; 96GB-00015730.

PA (MED1-) MEDICAL RES COUNCIL.

PI Maresden HS, Stow ND, McLean GW;

DR WPI; 1998-130695/12.

PT Antiviral agent capable of inhibiting herpes virus replication - disrupts
association between herpes simplex virus type 1 UL8 and POL.

PS Claim 5; Page 57; 83bp; English.

CC The present sequence represents an antiviral agent peptide capable of
combating herpes virus replication. The antiviral agent is capable of
disrupting the association between UL8 and POL (UL30), where UL8 and POL
are respectively defined as UL8 and POL of herpes simplex virus type 1
(HSV-1) together with homologues in other herpes viruses. The present
invention also describes an assay to determine the ability of a test
substance to interfere with the association of UL8 and POL. The assay
comprises: (a) exposing a 1st viral component to a test substance
followed by a 2nd viral component, or exposing a 1st viral component to a
2nd viral component followed by a test substance; (b) washing to remove
any 2nd viral component and/or test substance not associated with the 1st
viral component; and (c) detecting the presence, and optionally
determining the amount, of 2nd viral component associated with the 1st
viral component. The antiviral agent can be used to combat herpes virus
replication or infection

XX Sequence 26 AA;

Query Match 64.2%; Score 97; DB 2; Length 26;
Best Local Similarity 95.0%; Pred. No. 1.9e-08;
Matches 19; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 DEWVRSLLAVDAQHAKRVAS 20
1 DEWVRSLLAVDAQHAKRVAS 20

Db 7 DEWVRSLLAVDAQHAKRVAS 26

RESULT 3
AAU29536
ID AAU29536 standard; protein; 873 AA.
XX
AC AAU29536;

DT 18-DEC-2001 (first entry)

DE Novel human secreted protein #27.

XX Human; vaccination; gene therapy; nutritional supplement;

KW stem cell proliferation; haematopoiesis; nerve tissue regeneration;

XX immune suppression; immune stimulation; anti-inflammatory; leukaemia.

OS Homo sapiens.

PN W0200179449-A2.

PD 25-OCT-2001.

PF 16-APR-2001; 2001WO-US008656.

XX 18-APR-2000; 2000US-00552929.

PR 26-JAN-2001; 2001US-00770160.

PA (HYSE-) HYSEQ INC.

PI Tang YT, Liu C, Drmanac RT;

DR WPI; 2001-611725/70.

PT Nucleic acids encoding a range of human polypeptides, useful in genetic
vaccination, testing and therapy.

PS Claim 20; Page 162-163; 765bp; English.

CC The invention relates to novel human secreted polypeptides. The
CC polypeptides and antibodies to the polypeptides are useful for
CC determining the presence of or predisposition to a disease associated
CC with altered levels of polypeptide. The polypeptides are also useful for
CC identifying agents (agonists and antagonists) that bind to them. Cells
CC expressing the proteins are useful for identifying a therapeutic agent
CC for use in treatment of a pathology related to aberrant expression or
CC physiological interactions of the polypeptide. Vectors comprising the
CC nucleic acids encoding the polypeptides and cells genetically engineered
CC to express them are also useful for producing the proteins. The proteins
CC are useful in genetic vaccination, testing and therapy, and can be used
CC as nutritional supplements. They may be used to increase stem cell
CC proliferation, to regulate haematopoiesis, and in bone, cartilage, tendon
CC and/or nerve tissue growth or regeneration, immune suppression and/or
CC stimulation; as anti-inflammatory agents; and in treatment of leukaemias.
CC AAU9510-AAU3304 represent the amino acid sequences of novel human
CC secreted proteins of the invention

XX Sequence 873 AA;

Query Match 34.8%; Score 52.5; DB 4; Length 873;
Best Local Similarity 38.7%; Pred. No. 21;
Matches 12; Conservative 8; Mismatches 8; Indels 3; Gaps 2;

QY 1 DEWVRSLLAV--DAQHAKRVASGGLRFRRLN 29
1 DEWVRSLLAV--DAQHAKRVASGGLRFRRLN 29

RESULT 4
AAV60589
ID AAV60589 standard; protein; 221 AA.

XX AAV60589;

AC	AAW93276;	
AD	27-AUG-1999	(first entry)
AE		
AF		
AG		
AH		
AI		
AJ		
AK		
AL		
AM		
AN		
AO		
AP		
AQ		
AR		
AS		
AT		
AX		
AY		
AZ		
BA		
BB		
BC		
BD		
BE		
BF		
BG		
BH		
BI		
BJ		
BK		
BL		
BM		
BN		
BO		
BP		
BQ		
BR		
BS		
BT		
BU		
BV		
BW		
BX		
BY		
BZ		
CA		
CB		
CC		
CD		
CE		
CF		
CG		
CH		
CI		
CJ		
CK		
CL		
CM		
CN		
CO		
CP		
CQ		
CR		
CS		
CT		
CU		
CV		
CW		
CX		
CY		
CZ		
DA		
DB		
DC		
DD		
DE		
DF		
DG		
DH		
DI		
DJ		
DK		
DL		
DM		
DN		
DO		
DP		
DQ		
DR		
DS		
DT		
DU		
DV		
DW		
DX		
DY		
DZ		
EA		
EB		
EC		
ED		
EE		
EF		
EG		
EH		
EI		
EJ		
EK		
EL		
EM		
EN		
EO		
EP		
EQ		
ER		
ES		
ET		
EU		
EV		
EW		
EX		
EY		
EZ		
FA		
FB		
FC		
FD		
FE		
FF		
FG		
FH		
FI		
FJ		
FK		
FL		
FM		
FN		
FO		
FP		
FQ		
FR		
FS		
FT		
FU		
FV		
FW		
FX		
FY		
FZ		
GA		
GB		
GC		
GD		
GE		
GF		
GG		
GH		
GI		
GJ		
GK		
GL		
GM		
GN		
GO		
GP		
GQ		
GR		
GS		
GT		
GU		
GV		
GW		
GX		
GY		
GZ		
HA		
HB		
HC		
HD		
HE		
HF		
HG		
HH		
HI		
HJ		
HK		
HL		
HM		
HN		
HO		
HP		
HQ		
HR		
HS		
HT		
HU		
HV		
HW		
HX		
HY		
HZ		
IA		

AAV39367
ID AAV39367 standard; protein; 389 AA.
XX
AC AAV39367;
XX
DT 20-DEC-1999 (first entry)
XX
DE SBHWSBI cytokine signalling suppressor amino acid sequence.
XX
KW SBHWSBI; WD40 SOCS box; negative regulator; cytokine signal; cancer;
KW Janus kinase; Jak; obesity; inflammatory disorder; heart disease;
KW Crohn's disease; neuropathy; immune disorder.
XX
OS Homo sapiens.
XX
PN MO9949032-A1.
XX
PD 30-SEP-1999.
XX
PF 02-DEC-1998; 98MO-EP007806.
XX
PR 23-MAR-1998; 98GB-00006222.
PR 17-SEP-1998; 98GB-00020299.
XX
PA (SMIK) SMITHKLINE BEECHAM PLC.
XX
PI Michalovich D, Sims MA, Shaikh N;
XX WPI; 1999-591089/50.
DR N-PSDB; AA228319.
XX
PT A new polypeptide, designated SBHWSBI.
XX
PS Claim 3; Page 22-23; 41pp; English.
XX
CC This is the SBHWSBI polypeptide sequence. The SBHWSBI polypeptide is
CC believed to be a member of the WD40 SOCS box family. SOCS box proteins
CC are involved in acting as negative regulators of cytokine signalling, by
CC inhibiting the function of the Janus kinase (Jak) family of protein
CC kinases. The polypeptide and polynucleotide sequences of SBHWSBI can be
CC used to create antibodies against SBHWSBI, and used in a method for
CC screening to identify compounds which stimulate or inhibit the function
CC of the polypeptide. Molecules of the invention are useful to treat
CC diseases or disorders including cancer, obesity, inflammatory disorders,
CC heart disease, Crohn's disease, neuropathies, and immune disorders
XX
SQ Sequence 389 AA;
XX
Query Match 33.8%; Score 51; DB 2; Length 389;
Best Local Similarity 35.5%; Pred. No. 14;
Matches 11; Conservative 7; Mismatches 11; Indels 2; Gaps 1;
XX
QY 1 DEMWRSILAV--DAQHAKRVASEGLRFFRLN 29
DB 311 DRWRSVSPSHDGLHVASLADDDGVRFWRID 341
XX
RESULT 9
AAV39368
ID AAV39368 standard; protein; 421 AA.
XX
AC AAV39368;
XX
DT 20-DEC-1999 (first entry)
XX
DE SBHWSBI cytokine signalling suppressor polypeptide sequence #2.
XX
KW SBHWSBI; WD40 SOCS box; negative regulator; cytokine signal; cancer;
KW Janus kinase; Jak; obesity; inflammatory disorder; heart disease;
KW Crohn's disease; neuropathy; immune disorder.
XX
OS Homo sapiens.
XX

PN MO9949032-A1.
XX
PD 30-SEP-1999.
XX
PF 02-DEC-1998; 98MO-EP007806.
XX
PR 23-MAR-1998; 98GB-00006222.
PR 17-SEP-1998; 98GB-00020299.
XX
PA (SMIK) SMITHKLINE BEECHAM PLC.
XX
PI Michalovich D, Sims MA, Shaikh N;
XX WPI; 1999-591089/50.
DR N-PSDB; AA228320.
XX
PT A new polypeptide, designated SBHWSBI.
XX
PS Claim 19; Page 24; 41pp; English.
XX
CC This is an SBHWSBI polypeptide sequence. The nucleotide sequence was
CC derived from Expressed Sequence Tag (ESTs) sequences, prior to the
CC identification of the complete SBHWSBI polynucleotide sequence
CC (AA228319). The SBHWSBI polypeptides (AAV39367-139368) are believed to be
CC members of the WD40 SOCS box family. SOCS box proteins are involved in
CC acting as negative regulators of cytokine signalling, by inhibiting the
CC function of the Janus kinase (Jak) family of protein kinases. The
CC polypeptide and polynucleotide sequences of SBHWSBI can be used to create
CC antibodies against SBHWSBI, and used in a method for screening to
CC identify compounds which stimulate or inhibit the function of the
CC polypeptide. Molecules of the invention are useful to treat diseases or
CC disorders including cancer, obesity, inflammatory disorders, heart
CC disease, Crohn's disease, neuropathies, and immune disorders
XX
SQ Sequence 421 AA;
XX
Query Match 33.8%; Score 51; DB 2; Length 421;
Best Local Similarity 35.5%; Pred. No. 15;
Matches 11; Conservative 7; Mismatches 11; Indels 2; Gaps 1;
XX
QY 1 DEMWRSILAV--DAQHAKRVASEGLRFFRLN 29
DB 311 DRWRSVSPSHDGLHVASLADDDGVRFWRID 341
XX
RESULT 10
AAB12445
ID AAB12445 standard; protein; 421 AA.
XX
AC AAB12445;
XX
DT 20-OCT-2000 (first entry)
XX
DE Human WSB1 protein SEQ ID NO:4.
XX
KW Human; WSB1.
XX
OS Homo sapiens.
XX
PN CN1249346-A.
XX
PD 05-APR-2000.
XX
PF 30-SEP-1998; 98CN-00121922.
XX
PR 30-SEP-1998; 98CN-00121922.
XX
PA (UYFU-) UNIV FUDAN.
XX
PI Yu L, Fu Q, Zhao Y;
XX WPI; 2000-400724/35.
DR N-PSDB; AAA60596.
XX

XX Preparation of human gene sequence and its encoded polypeptide.
 PT
 PS Claim 4; Page 15-16; 24pp; Chinese.
 XX
 CC The present sequence represents human WSBI. Human WSBI is a homologue of
 CC mouse WSBI
 CC
 SQ Sequence 421 AA;

Query Match 33.8%; Score 51; DB 3; Length 421;
 Best Local Similarity 35.5%; Pred. No. 15;
 Matches 11; Conservative 7; Mismatches 11; Indels 2; Gaps 1;

QY 1 DEWVSLAV--DAQHAKRVASEGLRPFRLN 29
 DB 311 DRWVSVSFSHGLHVASLADDKVRFWRID 341

RESULT 11

AAV53885
 ID AAV53885 standard; protein; 421 AA.

AC AAV53885;

DT 13-MAR-2000 (first entry)

DE A suppressor of cytokine signalling protein designated HSCOP-5.

XX Human; suppressor of cytokine signalling protein; SOCS protein; HSCOP;
 KM cancer; leukaemia; lymphoma; diabetes mellitus; Crohn's disease;
 KM immune disorder; AIDS; allergy; atherosclerosis; inflammatory disorder;
 KM rheumatoid arthritis; irritable bowel syndrome; multiple sclerosis;
 KM ulcerative colitis; neurological disorder; Down's syndrome; amnesia;
 KM cerebral neoplasm; Huntington's disease; viral infection; adenovirus;
 KM acute respiratory disease; toga virus; rubella.

OS Homo sapiens.

XX Key Location/Qualifiers

FT Modified-site 18 /note= "potential phosphorylation site"
 FT Modified-site 37 /note= "potential glycosylation site"
 FT Modified-site 57 /note= "potential phosphorylation site"
 FT Modified-site 77 /note= "potential phosphorylation site"
 FT Modified-site 80 /note= "potential glycosylation site"
 FT Modified-site 83 /note= "potential glycosylation site"
 FT Modified-site 118 /note= "potential phosphorylation site"
 FT Modified-site 159 /note= "potential phosphorylation site"
 FT Modified-site 194 /note= "potential phosphorylation site"
 FT Modified-site 328 /note= "potential phosphorylation site"
 FT Modified-site 380 /note= "potential phosphorylation site"
 FT Modified-site 394 /note= "potential phosphorylation site"
 FT Modified-site 418 /note= "potential phosphorylation site"
 FT Modified-site 419 /note= "potential phosphorylation site"
 FT Modified-site /note= "potential phosphorylation site"

XX W09961614-A2.
 XX 02-DEC-1999.
 XX

PF 25-MAY-1999; 99WC-US011497.
 XX
 PR 28-MAY-1998; 98US-0087104P.
 PR 17-DEC-1998; 98US-0021600C.
 XX

XX (INCY-) INCYTE PHARM INC.

XX Ial F, Hillman JT, Gorgone G, Corley NC, Patterson C, Yue H;
 PI Tang YT, Azimzai Y;
 XX

DR WPI; 2000-072621/06.

DR N-PSDB; AA236829.

PT New purified polypeptide encoding human suppressor of cytokine signalling
 PT (SOCS) proteins useful for diagnosing, treating or preventing disorders
 PT associated with human SOCS proteins.

PS Claim 1; Page 76-77; 90pp; English.

CC The present sequence represents a human suppressor of cytokine signalling
 CC (SOCS) protein, designated HSCOP-5. The protein is useful for treating
 CC and/or preventing a disorder associated with decreased expression or
 CC activity of HSCOP. The protein antagonist is useful for treating and/or
 CC preventing a disorder associated with increased expression or activity of
 CC HSCOP. The human SOCS proteins and polynucleotides encoding them are
 CC useful in the diagnosis, treatment and prevention of cancer such as
 CC leukemia and lymphoma (especially e.g. cancers of the bone, heart and
 CC skin), diabetes mellitus, Crohn's disease, immune disorders e.g. AIDS,
 CC allergies and atherosclerosis, inflammatory disorders e.g. rheumatoid
 CC arthritis, irritable bowel syndrome, multiple sclerosis and ulcerative
 CC colitis, neurological disorders e.g. Down's syndrome, amnesia, cerebral
 CC neoplasms and Huntington's disease and infectious diseases such as those
 CC caused by viral infection e.g. adenoviruses (acute respiratory disease)
 CC and toga viruses (rubella) as well as those caused by bacterial, fungal,
 CC parasitic, protozoal and helminthic infections

XX Sequence 421 AA;

Query Match 33.8%; Score 51; DB 3; Length 421;
 Best Local Similarity 35.5%; Pred. No. 15;
 Matches 11; Conservative 7; Mismatches 11; Indels 2; Gaps 1;

QY 1 DEWVSLAV--DAQHAKRVASEGLRPFRLN 29
 DB 311 DRWVSVSFSHGLHVASLADDKVRFWRID 341

RESULT 12

AAV53410
 ID AAV53410 standard; protein; 421 AA.

AC AAV53410;

DT 06-NOV-2001 (first entry)

DE Human polypeptide, SEQ ID NO: 3020.

XX Human; full length cDNA; cDNA synthesis; oligo-capping.

OS Homo sapiens.

XX EP1130094-A2.

XX 05-SEP-2001.

PF 07-JUL-2000; 2000EP-00114089.

PR 08-JUL-1999; 99JP-00194486.

PR 11-JAN-2000; 2000JP-00118774.

XX 02-MAY-2000; 2000JP-00183765.
 XX (HELI-) HELIX RES INST.
 XX

CC The invention relates to primers for synthesising full length cDNA
CC clones. 330 cDNA molecules encoding a human protein have been isolated
CC and nucleotide sequences of 5'- and 3'-ends of the cDNA molecules have
CC been determined. Primers for synthesising the full length cDNA are used
CC for clarifying the function of the protein encoded by the cDNA. The full
CC length clones were obtained by construction of full length enriched cDNA
CC libraries that were synthesised by the oligo-capping method. The primers
CC enable the production of the full length cDNA easily without any special
CC methods. The present sequence is a polypeptide encoded by a full length
CC human cDNA of the invention. Note: The sequence data for this patent did
CC not form part of the printed specification, but was obtained in CD-ROM
CC format directly from EPO
XX
XX Sequence 421 AA;

```

02      1 DEMWRSIAY--DAQHAARKVASEGLRFFRLN 29
      ||||| | | : : : :
Db      311 DRWRSVSFSDGLHVASLADDKMVRFRWID 3411

```

```
RESULT 13
AAM39188
ID AAM39188 standard; protein; 421 AA
```

DT 22-OCT-2001 (first entry)

DE	Human polypeptide SEQ ID NO 2333.
DE	Human polypeptide SEQ ID NO 2333.

Human; nootropic; immunosuppressant; cytostatic; gene therapy; cancer; peripheral nervous system; neuropathy; central nervous system; CNS; Alzheimer's; Parkinson's disease; Huntington's disease; haemostatic; amyotrophic lateral sclerosis; Shy-Drager Syndrome; chemotactic; chemokinetic; thrombolytic; drug screening; arthritis; inflammation; leukaemia.

OS Homo sapiens.

PN WO200153312-A1.

PD 26-JUL-2001

PF 26-DEC-2000; 2000WO-US034263.

PR 23-DEC-1999; 99US-00471275.

PR 25-APR-2000; 2000US-00552317.

PR 19-JUL-2000; 2000US-00620312

PR 14-SEP-2000; 2000US-00662191

PR 29-NOV-2000; 2000US-00727344
XX

PA (HYSE-) HYSEQ INC.
XY

PI Tang YT, Liu C, Asundi V, Chen R, Ma Y, Qian XB, Ren F, Wang D,

PT Novel nucleic acids and polypeptides, useful for treating disorders such as central nervous system injuries.

PS Example 4; SEQ ID NO 2333; 10078pp; English

The invention relates to human nucleic acids (AA157798-AA161369) and the encoded polypeptides (AA038642-AA042213) with neurotropic, immunosuppressant and cytostatic activity. The polynucleotides are useful in gene therapy. A composition containing a polypeptide or polynucleotide of the invention may be used to treat diseases of the peripheral nervous system, such as peripheral nervous injuries, peripheral neuropathy and localised neuropathies and central nervous system diseases, such as Alzheimer's, Parkinson's disease, Huntington's disease, amyotrophic lateral sclerosis, and Shy-Drager Syndrome. Other uses include the utilisation of the activities such as: immune system suppression, Activin/inhibin activity, chemotactic/chemokinetic activity, haemostatic and thrombolytic activity, cancer diagnosis and therapy, drug screening, assays for receptor activity, arthritis and inflammation, leukemias and C.N.S disorders. Note: The sequence data for this patent did not form part of the printed specification

Sequence 421 AA;

Query Match	33.8%	Score 51	DB 4	Length 421
Best Local Similarity	35.5%	Pred. No. 15		
Matches 11, Conservative	7	Mismatches 11	Indels 2	Gaps 1

```
Qy      1 DEWTRSLAV--DAQHAAKRVASSEGLRFFRLN 29
          ||||::| | :|:::
Db      311 DRWRVSVFSHSDGLHVASLADDKMVRFWRID 341
```

RESULT 14
AAU82968
ID AAU82968 standard; protein; 611 AA

AC AAU82968;

DT 23-APR-2002 (first entry)

DE Human homologue of SQT1 protein target for antifungal compound.

KW Antifungal; fungal gene transcription; RPP34; POP3; TFA2; NAB2; MPT1;
 KW MTR2; BOS1; POL30; RSA2; SQT1; MTM1; TBP1; SPC98; BFR2; RNA1; GCD7; SKI6;
 KW NIP1; LGS5; NEB103; ECO1; ORC; CEN1; YPD1; TIM10; SRB4; yeast; fungus.

OS Homo sapiens.

PN WO200202055-A2

PD 10-JAN-2002
yy

28-JUN-2001; 2001WO-US020592.
PF
XX

PR 29-JUN-2000; 2000US-0215164P
PR 10-AUG-2000; 2000US-0234457P

XX
DA (ANAD-) ANADYS PHARM INC.

XX
PT Moore T. Buurman ET. Desilya T. Harris S. Komarnitsky S;

PT Davidoy E. Thompson CM:
 P1 Mendillo M, Moore D, McCoy M, Sandelson N, May J, Liu C, Wong J
 P2

XX
DR
WPI: 2002-147962/19

LN-FOUO; ABNC2000.
XX
XX

PT Screening candidate antitubercular compound for interaction with essential

Fri Oct 1 12:26:21 2004

us-09-988-851a-7.rapb

Page 1

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: September 30, 2004, 11:08:46 ; Search time 50.2857 Seconds
(without alignments)
127.988 Million cell updates/sec

Title: US-09-988-851a-7
Perfect score: 105
Sequence: 1 IEIVFTGVLAVGWGEGGKRV 20

Scoring table: BIOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 1351062 seqs, 32179191 residues

Total number of hits satisfying chosen parameters: 1351062

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :

Published Applications AA: *
1: /cgn2_6/ptodata/1/pubpaa/US07_PUBCOMB.pep:*
2: /cgn2_6/ptodata/1/pubpaa/PCT_NEW_PUB.pep:*
3: /cgn2_6/ptodata/1/pubpaa/US06_NEW_PUB.pep:*
4: /cgn2_6/ptodata/1/pubpaa/US06_PUBCOMB.pep:*
5: /cgn2_6/ptodata/1/pubpaa/US07_NEW_PUB.pep:*
6: /cgn2_6/ptodata/1/pubpaa/PCTUS_PUBCOMB.pep:*
7: /cgn2_6/ptodata/1/pubpaa/US08_NEW_PUB.pep:*
8: /cgn2_6/ptodata/1/pubpaa/US08_PUBCOMB.pep:*
9: /cgn2_6/ptodata/1/pubpaa/US09A_PUBCOMB.pep:*
10: /cgn2_6/ptodata/1/pubpaa/US09B_PUBCOMB.pep:*
11: /cgn2_6/ptodata/1/pubpaa/US09C_PUBCOMB.pep:*
12: /cgn2_6/ptodata/1/pubpaa/US09_NEW_PUB.pep:*
13: /cgn2_6/ptodata/1/pubpaa/US10A_PUBCOMB.pep:*
14: /cgn2_6/ptodata/1/pubpaa/US10B_PUBCOMB.pep:*
15: /cgn2_6/ptodata/1/pubpaa/US10C_PUBCOMB.pep:*
16: /cgn2_6/ptodata/1/pubpaa/US10_NEW_PUB.pep:*
17: /cgn2_6/ptodata/1/pubpaa/US60_NEW_PUB.pep:*
18: /cgn2_6/ptodata/1/pubpaa/US60_PUBCOMB.pep:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	105	100.0	20	12	US-09-998-851-7 Sequence 7, Appli
2	92	87.6	29	12	US-09-998-851-5 Sequence 5, Appli
3	82	78.1	15	12	US-09-998-851-6 Sequence 6, Appli
4	82	78.1	27	12	US-09-998-851-4 Sequence 4, Appli
5	71	67.6	25	12	US-09-998-851-3 Sequence 3, Appli
6	63	60.0	23	12	US-09-998-851-2 Sequence 2, Appli
7	52	49.5	25	10	US-09-988-067B-10 Sequence 10, Appli
8	52	49.5	25	12	US-10-335-977-6137 Sequence 6137, Ap
9	52	49.5	26	12	US-10-335-977-6138 Sequence 6138, Ap
10	48	45.7	3816	10	US-09-808-880-3 Sequence 3, Appli
11	47	44.8	85	12	US-10-424-599-228380 Sequence 228380,
12	47	44.8	238	12	US-10-425-114-53038 Sequence 53038, A
13	47	44.8	263	14	US-10-156-761-95732 Sequence 9572, Ap
14	47	44.8	482	16	US-10-767-701-46532 Sequence 46532, A
15	47	44.8	485	12	US-10-425-114-57489 Sequence 57489, A

16	46	43.8	15	12	US-09-998-851-8 Sequence 8, Appli
17	46	43.8	81	16	US-10-437-963-118743 Sequence 118743,
18	46	43.8	113	16	US-10-437-963-125727 Sequence 125727,
19	46	43.8	113	16	US-10-767-701-41883 Sequence 41883, A
20	46	43.8	225	12	US-10-424-599-208000 Sequence 208000,
21	46	43.8	276	12	US-10-424-599-235403 Sequence 235403,
22	46	43.8	3519	10	US-09-808-880-4 Sequence 4, Appli
23	46	43.8	4150	10	US-09-808-880-2 Sequence 2, Appli
24	46	43.8	7068	16	US-10-203-295-20 Sequence 20, Appli
25	46	43.8	9477	12	US-10-203-295-37 Sequence 37, Appli
26	45	42.9	223	12	US-10-335-977-5998 Sequence 5998, Ap
27	45	42.9	277	12	US-09-988-067B-82 Sequence 82, Appli
28	45	42.9	277	12	US-10-335-977-5999 Sequence 5999, Ap
29	45	42.9	373	15	US-10-074-978A-250 Sequence 250, App
30	45	42.9	388	12	US-10-425-114-54454 Sequence 54454, A
31	45	42.9	395	15	US-10-190-115-100 Sequence 100, App
32	45	42.9	395	15	US-10-369-072-100 Sequence 100, App
33	44.5	42.4	65	11	US-09-864-408A-4178 Sequence 4178, Ap
34	44	41.9	129	15	US-10-012-697-1515 Sequence 1515, Ap
35	44	41.9	139	14	US-10-148-759-72 Sequence 72, Appli
36	44	41.9	216	16	US-10-437-963-103994 Sequence 103994,
37	44	41.9	219	16	US-10-767-701-37174 Sequence 37174, A
38	44	41.9	255	14	US-10-156-761-14745 Sequence 14745, A
39	44	41.9	270	14	US-10-017-161-2266 Sequence 2266, Ap
40	44	41.9	270	15	US-10-292-798-1912 Sequence 1912, Ap
41	44	41.9	383	12	US-10-424-599-154558 Sequence 154558,
42	44	41.9	405	12	US-10-282-122A-60023 Sequence 60023, A
43	44	41.9	427	12	US-10-282-122A-43406 Sequence 43406, A
44	44	41.9	1612	16	US-10-647-156-26 Sequence 26, Appli
45	44	41.9	1620	16	US-10-437-963-192062 Sequence 192062,

ALIGNMENTS

RESULT 1
US-09-998-851-7
Sequence 7, Application US/09998851
Publication No. US20020034609A1
GENERAL INFORMATION:
APPLICANT: Medical Research Council
APPLICANT: Marsden, Howard
APPLICANT: Mcleat, Nigel
TITLE OR INVENTION: ASSAY FOR SCREENING FOR AN ANTI-VIRAL AGENT
FILE REFERENCE: P93472A
CURRENT APPLICATION NUMBER: US/09/999,851
CURRENT FILING DATE: 2002-03-11
NUMBER OF SEQ ID NOS: 18
SOFTWARE: PatentIn version 3.1
SEQ ID NO 7
LENGTH: 20
TYPE: PRT
ORGANISM: Artificial Sequence
FEATURE:
OTHER INFORMATION: Peptide derived from herpes simplex virus.
US-09-998-851-7

Query Match 100.0%; Score 105; DB 12; Length 20;
Best local similarity 100.0%; Pred. No. 2, 8e-08;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Cy 1 IEIVFTGVLAVGWGEGGKRV 20
Db 1 IEIVFTGVLAVGWGEGGKRV 20

RESULT 2
US-09-998-851-5
Sequence 5, Application US/09998851
Publication No. US20020034609A1
GENERAL INFORMATION:
APPLICANT: Medical Research Council

APPLICANT: Marsden, Howard
APPLICANT: Stow, Nigel
APPLICANT: McLean, Gordon
TITLE OF INVENTION: ASSAY FOR SCREENING FOR AN ANTI-VIRAL AGENT
FILE REFERENCE: P29347A
CURRENT APPLICATION NUMBER: US/09/998,851
CURRENT FILING DATE: 2002-03-11
NUMBER OF SEQ ID NOS: 18
SOFTWARE: PatentIn version 3.1
SEQ ID NO 5
LENGTH: 29
TYPE: PRT
ORGANISM: Artificial Sequence
FEATURE:
OTHER INFORMATION: Peptide derived from herpes simplex virus.
US-09-998-851-5

Query Match 87.6%; Score 92; DB 12; Length 29;
Best Local Similarity 100.0%; Pred. No. 3.1e-06;
Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 4 VTGVLGVWGGGKRV 20
DB 1 VTGVLGVWGGGKRV 17

RESULT 3
US-09-998-851-6
Sequence 6, Application US/09998851
Publication No. US20020034609A1
GENERAL INFORMATION:
APPLICANT: Medical Research Council
APPLICANT: Marsden, Howard
APPLICANT: Stow, Nigel
APPLICANT: McLean, Gordon
TITLE OF INVENTION: ASSAY FOR SCREENING FOR AN ANTI-VIRAL AGENT
FILE REFERENCE: P29347A
CURRENT APPLICATION NUMBER: US/09/998,851
CURRENT FILING DATE: 2002-03-11
NUMBER OF SEQ ID NOS: 18
SOFTWARE: PatentIn version 3.1
SEQ ID NO 6
LENGTH: 15
TYPE: PRT
ORGANISM: Artificial Sequence
FEATURE:
OTHER INFORMATION: Peptide derived from herpes simplex virus.
US-09-998-851-6

Query Match 78.1%; Score 82; DB 12; Length 15;
Best Local Similarity 100.0%; Pred. No. 4.4e-05;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 6 TGVLGVWGGGKRV 20
DB 1 TGVLGVWGGGKRV 15

RESULT 4
US-09-998-851-4
Sequence 4, Application US/09998851
Publication No. US20020034609A1
GENERAL INFORMATION:
APPLICANT: Medical Research Council
APPLICANT: Marsden, Howard
APPLICANT: Stow, Nigel
APPLICANT: McLean, Gordon
TITLE OF INVENTION: ASSAY FOR SCREENING FOR AN ANTI-VIRAL AGENT
FILE REFERENCE: P29347A
CURRENT APPLICATION NUMBER: US/09/998,851
CURRENT FILING DATE: 2002-03-11
NUMBER OF SEQ ID NOS: 18
SOFTWARE: PatentIn version 3.1

SEQ ID NO 4
LENGTH: 27
TYPE: PRT
ORGANISM: Artificial Sequence
FEATURE:
OTHER INFORMATION: Peptide derived from herpes simplex virus.
US-09-998-851-4

Query Match 78.1%; Score 82; DB 12; Length 27;
Best Local Similarity 100.0%; Pred. No. 8e-05;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 6 TGVLGVWGGGKRV 20
DB 1 TGVLGVWGGGKRV 15

RESULT 5
US-09-998-851-3
Sequence 3, Application US/09998851
Publication No. US20020034609A1
GENERAL INFORMATION:
APPLICANT: Medical Research Council
APPLICANT: Marsden, Howard
APPLICANT: Stow, Nigel
APPLICANT: McLean, Gordon
TITLE OF INVENTION: ASSAY FOR SCREENING FOR AN ANTI-VIRAL AGENT
FILE REFERENCE: P29347A
CURRENT APPLICATION NUMBER: US/09/998,851
CURRENT FILING DATE: 2002-03-11
NUMBER OF SEQ ID NOS: 18
SOFTWARE: PatentIn version 3.1
SEQ ID NO 3
LENGTH: 25
TYPE: PRT
ORGANISM: Artificial Sequence
FEATURE:
OTHER INFORMATION: Peptide derived from herpes simplex virus.
US-09-998-851-3

Query Match 67.6%; Score 71; DB 12; Length 25;
Best Local Similarity 100.0%; Pred. No. 0.0028;
Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 8 VLVGVWGGGKRV 20
DB 1 VLVGVWGGGKRV 13

RESULT 6
US-09-998-851-2
Sequence 2, Application US/09998851
Publication No. US20020034609A1
GENERAL INFORMATION:
APPLICANT: Medical Research Council
APPLICANT: Marsden, Howard
APPLICANT: Stow, Nigel
APPLICANT: McLean, Gordon
TITLE OF INVENTION: ASSAY FOR SCREENING FOR AN ANTI-VIRAL AGENT
FILE REFERENCE: P29347A
CURRENT APPLICATION NUMBER: US/09/998,851
CURRENT FILING DATE: 2002-03-11
NUMBER OF SEQ ID NOS: 18
SOFTWARE: PatentIn version 3.1
SEQ ID NO 2
LENGTH: 23
TYPE: PRT
ORGANISM: Artificial Sequence
FEATURE:
OTHER INFORMATION: Peptide derived from herpes simplex virus.
US-09-998-851-2

Query Match 60.0%; Score 63; DB 12; Length 23;

Best Local Similarity 100.0%; Pred. No. 0.037;
Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 10 AGVWGGGKRV 20
Db 1 AGVWGGGKRV 11

RESULT 7

US-09-988-067B-10
; Sequence 10, Application US/09988067B
; Publication No. US20030124141A1
; GENERAL INFORMATION:
; APPLICANT: Haas, Rainer
; APPLICANT: Kleantous, Harold
; APPLICANT: Tomb, Jean-Francois
; APPLICANT: Miller, Charles
; APPLICANT: Al-Garawi, Amal
; APPLICANT: Odenbreit, Stefan
; APPLICANT: Meyer, Thomas
; TITLE OF INVENTION: Helicobacter Polypeptides and
; TITLE OF INVENTION: Corresponding Polynucleotide Molecules
; FILE REFERENCE: 06132/040002
; CURRENT APPLICATION NUMBER: US/09/988,067B
; CURRENT FILING DATE: 2003-01-31
; PRIOR APPLICATION NUMBER: US 08/831,309
; PRIOR FILING DATE: 1997-04-01
; NUMBER OF SEQ ID NOS: 112
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 10
; LENGTH: 255
; TYPE: PRT
; ORGANISM: Helicobacter pylori
US-09-988-067B-10

Query Match 49.5%; Score 52; DB 10; Length 255;
Best Local Similarity 47.1%; Pred. No. 16;
Matches 8; Conservative 6; Mismatches 3; Indels 0; Gaps 0;

QY 4 VFTGVLAGVWGGGKRV 20
Db 12 LFTGILSGIFGIGGLI 28

RESULT 8

US-10-335-977-6137
; Sequence 6137, Application US/10335977
; Publication No. US20040052799A1
; GENERAL INFORMATION:
; APPLICANT: DOUGLAS SMITH et al
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES
; RELATING TO HELICOBACTER PYLORI FOR
; DIAGNOSTICS AND THERAPEUTICS
; NUMBER OF SEQUENCES: 10031
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: LAHIVE & COCKFIELD
; STREET: 28 State Street
; CITY: Boston
; STATE: Massachusetts
; COUNTRY: USA
; ZIP: 02109-1875
; COMPUTER READABLE FORM:
; MEDIUM TYPE: CD-ROM ISO9660
; COMPUTER: IBM PC Compatible
; OPERATING SYSTEM: Windows NT 4.0
; SOFTWARE: UNIX
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/10/335,977
; FILING DATE: 30-Dec-2002
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/993,002
; FILING DATE: 17-DEC-1997
; ATTORNEY/AGENT INFORMATION:

NAME: Mandragouras, Amy E.
REGISTRATION NUMBER: 36,207
REFERENCE/DOCKET NUMBER: GTN-018
TELECOMMUNICATION INFORMATION:
TELEPHONE: (617)227-7400
TELEFAX: (617)742-4214
INFORMATION FOR SEQ ID NO: 6137:
SEQUENCE CHARACTERISTICS:
LENGTH: 255 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
HYPOTHETICAL: YES
ORIGINAL SOURCE:
ORGANISM: Helicobacter pylori
FEATURE:
NAME/KEY: misc.feature
LOCATION: (B) LOCATION 1...255
SEQUENCE DESCRIPTION: SEQ ID NO: 6137:
US-10-335-977-6137

Query Match 49.5%; Score 52; DB 12; Length 255;
Best Local Similarity 47.1%; Pred. No. 16;
Matches 8; Conservative 6; Mismatches 3; Indels 0; Gaps 0;

QY 4 VFTGVLAGVWGGGKRV 20
Db 12 LFTGILSGIFGIGGLI 28

RESULT 9

US-10-335-977-6138
; Sequence 6138, Application US/10335977
; Publication No. US20040052799A1
; GENERAL INFORMATION:
; APPLICANT: DOUGLAS SMITH et al
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES
; RELATING TO HELICOBACTER PYLORI FOR
; DIAGNOSTICS AND THERAPEUTICS
; NUMBER OF SEQUENCES: 10031
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: LAHIVE & COCKFIELD
; STREET: 28 State Street
; CITY: Boston
; STATE: Massachusetts
; COUNTRY: USA
; ZIP: 02109-1875
; COMPUTER READABLE FORM:
; MEDIUM TYPE: CD-ROM ISO9660
; COMPUTER: IBM PC Compatible
; OPERATING SYSTEM: Windows NT 4.0
; SOFTWARE: UNIX
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/10/335,977
; FILING DATE: 30-Dec-2002
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/993,002
; FILING DATE: 17-DEC-1997
; ATTORNEY/AGENT INFORMATION:
; NAME: Mandragouras, Amy E.
; REGISTRATION NUMBER: 36,207
; REFERENCE/DOCKET NUMBER: GTN-018
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (617)227-7400
; TELEFAX: (617)742-4214
; INFORMATION FOR SEQ ID NO: 6138:
SEQUENCE CHARACTERISTICS:
LENGTH: 264 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
HYPOTHETICAL: YES
ORIGINAL SOURCE:

ORGANISM: Helicobacter pylori
FEATURE:
NAME/KEY: misc_feature
LOCATION: (B) LOCATION 1...264
SEQUENCE DESCRIPTION: SEQ ID NO: 6138
US-10-335-977-6:38

Query Match 49.5%; Score 52; DB 12; Length 264;
Best Local Similarity 47.1%; Pred. No. 17;
Matches 8; Conservative 6; Mismatches 3; Indels 0; Gaps 0;

QY 4 VFTGVLAGWGSQKRV 20
DB 21 LFTGILSGIFGIGGMI 37

RESULT 10
US-09-808-880-3
Sequence 3, Application US/09808880
Publication No. US2003002787A1
GENERAL INFORMATION:
APPLICANT: Betlach, Mary C.
APPLICANT: Shah, Sanjay Krishnakant
APPLICANT: McDaniel, Robert
APPLICANT: Tang, Li
TITLE OF INVENTION: RECOMBINANT OLEANDOLIDE POLYKETIDE SYNTHASE
FILE REFERENCE: 30062-20029.00
CURRENT APPLICATION NUMBER: US/09/808,880
CURRENT FILING DATE: 2001-03-14
PRIOR APPLICATION NUMBER: US/09/428,517
PRIOR FILING DATE: 1999-10-28
PRIOR APPLICATION NUMBER: 60/120,254
PRIOR FILING DATE: 1999-02-16
PRIOR APPLICATION NUMBER: 60/106,100
PRIOR FILING DATE: 1998-10-29
NUMBER OF SEQ ID NOS: 12
SOFTWARE: Patent In Ver. 2.1
SEQ ID NO 3
LENGTH: 3616
TYPE: PRT
ORGANISM: Artificial Sequence
FEATURE:
OTHER INFORMATION: Description of Artificial Sequence: Recombinant
US-09-808-880-3

Query Match 45.7%; Score 48; DB 10; Length 3616;
Best Local Similarity 47.1%; Pred. No. 9,4e+02;
Matches 8; Conservative 4; Mismatches 5; Indels 0; Gaps 0;

QY 2 ELVFTGVLAGWGSQK 18
DB 1341 DVEFSSVAGWGSAGQ 1357

RESULT 11
US-10-424-599-228380
Sequence 228380, Application US/10424599
Publication No. US20040031072A1
GENERAL INFORMATION:
APPLICANT: La Rosa Thomas J
APPLICANT: Kovalic David K
APPLICANT: Zhou Yihua
APPLICANT: Cao Yongwei
TITLE OF INVENTION: Soy Nucleic Acid Molecules and Other Molecules Associated With
FILE REFERENCE: 38-21(53323)B
CURRENT APPLICATION NUMBER: US/10/424,599
CURRENT FILING DATE: 2003-04-28
NUMBER OF SEQ ID NOS: 285684
SEQ ID NO 228380
LENGTH: 85
TYPE: PRT
ORGANISM: Streptomyces avermitilis

ORGANISM: Glycine max
FEATURE:
OTHER INFORMATION: Clone ID: PAT_MRT3847_48255C.1.pcp
US-10-424-599-228380

Query Match 44.8%; Score 47; DB 12; Length 85;
Best Local Similarity 61.5%; Pred. No. 28;
Matches 8; Conservative 1; Mismatches 4; Indels 0; Gaps 0;

QY 4 VFTGVLAGWGSQK 16
DB 33 IFGTRPAGFWGEG 45

RESULT 12
US-10-425-114-53038
Sequence 53038, Application US/10425114
Publication No. US20040034888A1
GENERAL INFORMATION:
APPLICANT: Liu, Jingdong
APPLICANT: Zhou, Yihua
APPLICANT: Kovalic, David K.
APPLICANT: Screen, Steven E
APPLICANT: Tabaska, Jack E
APPLICANT: Cao, Yongwei
TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
FILE REFERENCE: 38-21(53313)B
CURRENT APPLICATION NUMBER: US/10/425,114
CURRENT FILING DATE: 2003-04-28
NUMBER OF SEQ ID NOS: 73128
SEQ ID NO 53038
LENGTH: 238
TYPE: PRT
ORGANISM: Zea mays
FEATURE:
OTHER INFORMATION: Clone ID: LIB3060-006-FL1_FLI.pcp
US-10-425-114-53038

Query Match 44.8%; Score 47; DB 12; Length 238;
Best Local Similarity 53.8%; Pred. No. 80;
Matches 7; Conservative 3; Mismatches 3; Indels 0; Gaps 0;

QY 3 LVFTGVLAGWGE 15
DB 170 LIFISILAGFWGQ 182

RESULT 13
US-10-136-761-9572
Sequence 9572, Application US/10156761
Publication No. US20030119018A1
GENERAL INFORMATION:
APPLICANT: OKURA, SATOSHI
APPLICANT: IKEDA, HARUO
APPLICANT: ISHIKAWA, JUN
APPLICANT: HORIKAWA, HIROSHI
APPLICANT: SHIBA, TADAYOSHI
APPLICANT: SAKAKI, YOSHIYUKI
APPLICANT: HATTORI, MASAHISA
TITLE OF INVENTION: NOVEL POLYNUCLEOTIDES
FILE REFERENCE: 249-262
CURRENT APPLICATION NUMBER: US/10/156,761
CURRENT FILING DATE: 2002-05-29
PRIOR APPLICATION NUMBER: JP 2001-204089
PRIOR FILING DATE: 2001-05-30
PRIOR APPLICATION NUMBER: JP 2001-272697
PRIOR FILING DATE: 2001-08-02
NUMBER OF SEQ ID NOS: 15109
SEQ ID NO 9572
LENGTH: 263
TYPE: PRT
ORGANISM: Streptomyces avermitilis

US-10-156-761-9572

Query Match 44.8%; Score 47; DB 14; Length 263;
Best Local Similarity 62.5%; Pred. No. 88;
Matches 10; Conservative 2; Mismatches 2; Indels 2; Gaps 1;

OY 3 LVFTGVLAGV--MGEG 16
DB 222 LVFTGVLAGV--MGEG 237

RESULT 14

US-10-767-701-46532
; Sequence 46532, Application US/10767701
; Publication No. US20040172684A1
; GENERAL INFORMATION:
; APPLICANT: Kovalic, David K.
; APPLICANT: Zhou, Yihua
; APPLICANT: Cao, Yongwei
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
; FILE REFERENCE: 38-21(53535)B
; CURRENT APPLICATION NUMBER: US/10/767,701
; CURRENT FILING DATE: 2004-01-29
; NUMBER OF SEQ ID NOS: 63128
; SEQ ID NO 46532
; LENGTH: 482
; TYPE: PRT
; ORGANISM: Sorghum bicolor
; FEATURE:
; OTHER INFORMATION: Clone ID: SORBI-28MAY03-C41778_1.pep
US-10-767-701-46532

Query Match 44.8%; Score 47; DB 16; Length 482;
Best Local Similarity 53.8%; Pred. No. 1.6e+02;
Matches 7; Conservative 3; Mismatches 3; Indels 0; Gaps 0;

OY 3 LVFTGVLAGV--MGEG 15
DB 414 LVFTGVLAGV--MGEG 426

RESULT 15

US-10-425-114-57489
; Sequence 57489, Application US/10425114
; Publication No. US20040034888A1
; GENERAL INFORMATION:
; APPLICANT: Liu, Jingdong
; APPLICANT: Zhou, Yihua
; APPLICANT: Kovalic, David K.
; APPLICANT: Screen, Steven E
; APPLICANT: Tabaska, Jack E
; APPLICANT: Cao, Yongwei
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
; FILE REFERENCE: 38-21(53313)B
; CURRENT APPLICATION NUMBER: US/10/425,114
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 73128
; SEQ ID NO 57489
; LENGTH: 485
; TYPE: PRT
; ORGANISM: Zea mays
; FEATURE:
; OTHER INFORMATION: Clone ID: UC-ZMROB73024E12_F11.pep
US-10-425-114-57489

Query Match 44.8%; Score 47; DB 12; Length 485;
Best Local Similarity 53.8%; Pred. No. 1.6e+02;
Matches 7; Conservative 3; Mismatches 3; Indels 0; Gaps 0;

OY 3 LVFTGVLAGV--MGEG 15
DB 414 LVFTGVLAGV--MGEG 426

DB 417 LVFTGVLAGV--MGEG 429

Search completed: September 30, 2004, 11:33:34
Job time: 51.2857 secs

This Page Blank (uspto)

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: September 30, 2004, 10:58:43 : Search time 13.9048 Seconds
(without alignments)
74.257 Million cell updates/sec

Title: US-09-988-851A-7
Perfect score: 105
Sequence: 1 IEVFTGVLAGVWGEGRKFV 20

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 389414 seqs, 51625971 residues

Total number of hits satisfying chosen parameters: 389414

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Issued Patents AA:
1: /cgn2_6/ptodata/2/iaa/5A.COMB.pep:*
2: /cgn2_6/ptodata/2/iaa/5B.COMB.pep:*
3: /cgn2_6/ptodata/2/iaa/6A.COMB.pep:*
4: /cgn2_6/ptodata/2/iaa/6B.COMB.pep:*
5: /cgn2_6/ptodata/2/iaa/PCTus.COMB.pep:*
6: /cgn2_6/ptodata/2/iaa/backfilest.pep:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	105	100.0	20	4	US-09-230-405-7
2	92	87.6	29	4	US-09-230-405-5
3	82	78.1	15	4	US-09-230-405-6
4	82	78.1	27	4	US-09-230-405-4
5	71	67.6	25	4	US-09-230-405-3
6	63	60.0	23	4	US-09-230-405-2
7	48	45.7	3816	3	US-09-428-517-3
8	46	43.8	15	4	US-09-230-405-8
9	46	43.8	1580	2	US-08-804-227C-11
10	46	43.8	1580	2	US-08-804-198-5
11	46	43.8	1891	2	US-08-804-227C-12
12	46	43.8	1891	2	US-08-804-198-6
13	46	43.8	3519	3	US-09-428-517-4
14	46	43.8	4150	3	US-09-428-517-2
15	45	42.9	32	1	US-08-615-279-26
16	45	42.9	138	4	US-09-570-921-13
17	45	42.9	173	4	US-09-489-039A-7555
18	45	42.9	192	4	US-09-543-681A-5012
19	45	42.9	679	4	US-09-252-991A-1861
20	45	42.9	1238	4	US-09-252-991A-26363
21	44.5	42.4	4545	2	US-08-804-227C-14
22	44.5	42.4	4550	2	US-08-804-227C-8
23	44.5	42.4	4550	2	US-08-804-198-2
24	44	41.9	533	1	US-08-294-872-2
25	44	41.9	533	5	PCT-US95-09823-2
26	43	41.0	26	2	US-08-394-021-7
27	43	41.0	26	4	US-09-131-551-7

28	43	41.0	198	3	US-08-965-056-79	Sequence 79, Appl
29	43	41.0	266	4	US-09-328-352-7793	Sequence 7793, Ap
30	43	41.0	332	4	US-09-252-991A-1791	Sequence 31791, A
31	43	41.0	596	2	US-08-392-806A-2	Sequence 2, Appl
32	43	41.0	596	3	US-09-257-490-2	Sequence 2, Appl
33	43	41.0	6095	3	US-09-144-085-2	Sequence 2, Appl
34	42.5	40.5	652	1	US-08-261-663A-6	Sequence 6, Appl
35	42.5	40.5	652	4	US-09-357-206A-5	Sequence 5, Appl
36	42.5	40.5	652	4	US-08-813-742A-5	Sequence 5, Appl
37	42.5	40.5	652	5	PCT-US95-07754A-6	Sequence 108, App
38	42.5	40.5	1143	2	US-08-310-912A-108	Sequence 108, App
39	42.5	40.5	1143	3	US-09-301-085-108	Sequence 108, App
40	42.5	40.5	1143	5	PCT-US95-04589-108	Sequence 108, App
41	42.5	40.5	1144	1	US-08-261-663A-2	Sequence 2, Appl
42	42.5	40.5	1144	1	US-08-261-663A-4	Sequence 4, Appl
43	42.5	40.5	1144	4	US-08-930-996A-9	Sequence 9, Appl
44	42.5	40.5	1144	4	US-09-357-206A-3	Sequence 3, Appl
45	42.5	40.5	1144	4	US-09-813-742A-3	Sequence 3, Appl

ALIGNMENTS

RESULT 1
US-09-230-405-7
; Sequence 7, Application US/09230405
; Patent No. 6337074
; GENERAL INFORMATION:
; APPLICANT: Medical Research Council
; TITLE OF INVENTION: ANTI-VIRAL AGENT
; FILE REFERENCE: P17970C
; CURRENT APPLICATION NUMBER: US/09/230,405
; CURRENT FILING DATE: 1999-01-25
; NUMBER OF SEQ ID NOS: 18
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 7
; LENGTH: 20
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence:peptide derived
; OTHER INFORMATION: from herpes simplex virus
US-09-230-405-7

Query Match 100.0%; Score 105; DB 4; Length 20;
Best Local Similarity 100.0%; Pred. No. 1.5e-08;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 IEVFTGVLAGVWGEGRKFV 20
DB 1 IEVFTGVLAGVWGEGRKFV 20

RESULT 2
US-09-230-405-5
; Sequence 5, Application US/09230405
; Patent No. 6337074
; GENERAL INFORMATION:
; APPLICANT: Medical Research Council
; TITLE OF INVENTION: ANTI-VIRAL AGENT
; FILE REFERENCE: P17970C
; CURRENT APPLICATION NUMBER: US/09/230,405
; CURRENT FILING DATE: 1999-01-25
; NUMBER OF SEQ ID NOS: 18
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 5
; LENGTH: 29
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence:peptide derived
; OTHER INFORMATION: from herpes simplex virus
US-09-230-405-5

Query Match 87.6%; Score 92; DB 4; Length 29;
Best Local Similarity 100.0%; Pred. No. 1.5e-06;
Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 4 VFTGLAGWGEQKRV 20
Db 1 VFTGLAGWGEQKRV 17

RESULT 3
US-09-230-405-6
Sequence 6, Application US/09230405

Patent No. 6337074
GENERAL INFORMATION:
APPLICANT: Medical Research Council
TITLE OF INVENTION: ANTI-VIRAL AGENT
FILE REFERENCE: PI7970C
CURRENT APPLICATION NUMBER: US/09/230.405
CURRENT FILING DATE: 1999-01-25
NUMBER OF SEQ ID NOS: 18
SOFTWARE: Patentin Ver. 2.1
SEQ ID NO 6
LENGTH: 15
TYPE: PRT
ORGANISM: Artificial Sequence
FEATURE:
OTHER INFORMATION: Description of Artificial Sequence:peptide derived
US-09-230-405-6

Query Match 78.1%; Score 82; DB 4; Length 15;
Best Local Similarity 100.0%; Pred. No. 2e-05;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 6 TGVLAGWGEQKRV 20
Db 1 TGVLAGWGEQKRV 15

RESULT 4
US-09-230-405-4
Sequence 4, Application US/09230405
Patent No. 6337074
GENERAL INFORMATION:
APPLICANT: Medical Research Council
TITLE OF INVENTION: ANTI-VIRAL AGENT
FILE REFERENCE: PI7970C
CURRENT APPLICATION NUMBER: US/09/230.405
CURRENT FILING DATE: 1999-01-25
NUMBER OF SEQ ID NOS: 18
SOFTWARE: Patentin Ver. 2.1
SEQ ID NO 4
LENGTH: 27
TYPE: PRT
ORGANISM: Artificial Sequence
FEATURE:
OTHER INFORMATION: Description of Artificial Sequence:peptide derived
US-09-230-405-4

Query Match 78.1%; Score 82; DB 4; Length 27;
Best Local Similarity 100.0%; Pred. No. 3.6e-05;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 6 TGVLAGWGEQKRV 20
Db 1 TGVLAGWGEQKRV 15

RESULT 5
US-09-230-405-3
Sequence 3, Application US/09230405

Patent No. 6337074
GENERAL INFORMATION:
APPLICANT: Medical Research Council
TITLE OF INVENTION: ANTI-VIRAL AGENT
FILE REFERENCE: PI7970C
CURRENT APPLICATION NUMBER: US/09/230.405
CURRENT FILING DATE: 1999-01-25
NUMBER OF SEQ ID NOS: 18
SOFTWARE: Patentin Ver. 2.1
SEQ ID NO 3
LENGTH: 25
TYPE: PRT
ORGANISM: Artificial Sequence
FEATURE:
OTHER INFORMATION: Description of Artificial Sequence:peptide derived
US-09-230-405-3

Query Match 67.6%; Score 71; DB 4; Length 25;
Best Local Similarity 100.0%; Pred. No. 0.0012;
Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 8 VLAGWGEQKRV 20
Db 1 VLAGWGEQKRV 13

RESULT 6
US-09-230-405-2
Sequence 2, Application US/09230405
Patent No. 6337074
GENERAL INFORMATION:
APPLICANT: Medical Research Council
TITLE OF INVENTION: ANTI-VIRAL AGENT
FILE REFERENCE: PI7970C
CURRENT APPLICATION NUMBER: US/09/230.405
CURRENT FILING DATE: 1999-01-25
NUMBER OF SEQ ID NOS: 18
SOFTWARE: Patentin Ver. 2.1
SEQ ID NO 2
LENGTH: 23
TYPE: PRT
ORGANISM: Artificial Sequence
FEATURE:
OTHER INFORMATION: Description of Artificial Sequence:peptide derived
US-09-230-405-2

Query Match 60.0%; Score 63; DB 4; Length 23;
Best Local Similarity 100.0%; Pred. No. 0.015;
Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 10 AGWGEQKRV 20
Db 1 AGWGEQKRV 11

RESULT 7
US-09-428-517-3
Sequence 3, Application US/09428517
Patent No. 6251636
GENERAL INFORMATION:
APPLICANT: Betlach, Mary C.
APPLICANT: Shah, Sanjay Krishnakant
APPLICANT: McDaniel, Robert
APPLICANT: Tang, Li
TITLE OF INVENTION: RECOMBINANT OLEANDROLE POLYKETIDE SYNTHASE
FILE REFERENCE: 30062-20029.00
CURRENT APPLICATION NUMBER: US/09/428.517
CURRENT FILING DATE: 1999-10-28
EARLIER APPLICATION NUMBER: 60/120.254
EARLIER FILING DATE: 1999-02-16
EARLIER APPLICATION NUMBER: 60/106.100

EARLIER FILING DATE: 1998-10-29
NUMBER OF SEQ ID NOS: 12
SOFTWARE: Patent In Ver. 2.1
SEQ ID NO 3
LENGTH: 3816
TYPE: PRT
ORGANISM: Artificial Sequence
FEATURE:
OTHER INFORMATION: Description of Artificial Sequence: Recombinant
US-09-428-517-3

Query Match 45.7%; Score 48; DB 3; Length 3816;
Best Local Similarity 47.1%; Pred. No. 3.2e+02;
Matches 8; Conservative 4; Mismatches 5; Indels 0; Gaps 0;

QY 2 ELVFTGVLAGVWGSGK 18
Db 1341 DVFFSSVAGVWGSAGQ 1357

RESULT 8
US-09-230-405-8
Sequence 8, Application US/09230405
Patent No. 6337074
GENERAL INFORMATION:
APPLICANT: Medical Research Council
TITLE OF INVENTION: ANTI-VIRAL AGENT
FILE REFERENCE: P17970C
CURRENT APPLICATION NUMBER: US/09/230,405
CURRENT FILING DATE: 1999-01-25
NUMBER OF SEQ ID NOS: 18
SOFTWARE: Patent In Ver. 2.1
SEQ ID NO 8
LENGTH: 15
TYPE: PRT
ORGANISM: Artificial Sequence
FEATURE:
OTHER INFORMATION: Description of Artificial Sequence: peptide derived
US-09-230-405-8

Query Match 43.8%; Score 46; DB 4; Length 15;
Best Local Similarity 100.0%; Pred. No. 2.4;
Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 IELVFTGVLA 10
Db 6 IELVFTGVLA 15

RESULT 9
US-08-804-227C-11
Sequence 11, Application US/08804227C
Patent No. 5876991
GENERAL INFORMATION:
APPLICANT: Dehoff, Bradley S.
APPLICANT: Kuhnstoss, Stuart A.
APPLICANT: Rostock, Paul R., Jr.
APPLICANT: Sutton, Kimberly L.
TITLE OF INVENTION: POLYKETIDE SYNTHASE GENES
NUMBER OF SEQUENCES: 15
CORRESPONDENCE ADDRESS:
ADDRESSEE: THOMAS G. PLANT 1501
STREET: LILLY CORPORATE CENTER
CITY: INDIANAPOLIS
STATE: IN
COUNTRY: USA
ZIP: 46285
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM Compatible
OPERATING SYSTEM: MS-DOS

SOFTWARE: ASCII(DOS) Text only
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/804,227C
FILING DATE: February 21, 1997
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: Plant, Thomas, G.
REGISTRATION NUMBER: 35,784
REFERENCE/DOCKET NUMBER: X-8231
TELECOMMUNICATION INFORMATION:
TELEPHONE: 317-276-2459
INFORMATION FOR SEQ ID NO: 11:
SEQUENCE CHARACTERISTICS:
LENGTH: 1580 amino acids
TYPE: amino acid
TOPOLOGY: unknown
MOLECULE TYPE: peptide
US-08-804-227C-11

Query Match 43.8%; Score 46; DB 2; Length 1580;
Best Local Similarity 72.7%; Pred. No. 2.5e+02;
Matches 8; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 10 AGVWGGGKTV 20
Db 1287 AGVWGGGQGV 1297

RESULT 10
US-08-804-198-5
Sequence 5, Application US/08804198
Patent No. 5945320
GENERAL INFORMATION:
APPLICANT: Bargett, Stanley G.
APPLICANT: Kuhnstoss, Stuart A.
APPLICANT: Rao, Nagaraja R.
APPLICANT: Richardson, Mark A.
APPLICANT: Rostock, Paul R., Jr.
TITLE OF INVENTION: PLANTENOLIDE SYNTHASE GENE
NUMBER OF SEQUENCES: 6
CORRESPONDENCE ADDRESS:
ADDRESSEE: PAUL R. CANTRELL 1138
STREET: LILLY CORPORATE CENTER
CITY: INDIANAPOLIS
STATE: IN
COUNTRY: USA
ZIP: 46285
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: Macintosh
OPERATING SYSTEM: Macintosh 7.0
SOFTWARE: Microsoft Word 5.1
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/804,198
FILING DATE:
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: Cantrell, Paul R.
REGISTRATION NUMBER: 36,470
REFERENCE/DOCKET NUMBER: P9113
TELECOMMUNICATION INFORMATION:
TELEPHONE: 317-276-3885
INFORMATION FOR SEQ ID NO: 5:
SEQUENCE CHARACTERISTICS:
LENGTH: 1580 amino acids
TYPE: amino acid
TOPOLOGY: unknown
MOLECULE TYPE: peptide
US-08-804-198-5

Query Match 43.8%; Score 46; DB 2; Length 1580;
Best Local Similarity 72.7%; Pred. No. 2.5e+02;
Matches 8; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 10 AGWGGGKRV 20
Db 1287 AGWGGGQSV 1297

RESULT 11

US-08-804-227C-12
Sequence 12, Application US/08804227C
Patent No. 5876991

GENERAL INFORMATION:

APPLICANT: Dehoff, Bradley S.
APPLICANT: Kustoss, Stuart A.
APPLICANT: Rostock, Paul R., Jr.
APPLICANT: Sutton, Kimberly L.
TITLE OF INVENTION: POLYKETIDE SYNTHASE GENES
NUMBER OF SEQUENCES: 15
CORRESPONDENCE ADDRESS: 15
ADDRESSEE: THOMAS G. PLANT 1501
STREET: LILLY CORPORATE CENTER
CITY: INDIANAPOLIS
STATE: IN
COUNTRY: USA
ZIP: 46285

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk
COMPUTER: IBM Compatible
OPERATING SYSTEM: MS-DOS
SOFTWARE: ASCII(DOS) Text only
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/804,227C
FILING DATE: February 21, 1997
CLASSIFICATION: 435

ATTORNEY/AGENT INFORMATION:

NAME: Plant, Thomas, G.
REGISTRATION NUMBER: 35,784
REFERENCE/DOCKET NUMBER: X-8231
TELECOMMUNICATION INFORMATION:
TELEPHONE: 317-276-2459
INFORMATION FOR SEQ ID NO: 12:

SEQUENCE CHARACTERISTICS:

LENGTH: 1891 amino acids
TYPE: amino acid
TOPOLOGY: unknown
MOLECULE TYPE: peptide
US-08-804-227C-12

Query Match

Best Local Similarity 43.8%; Score 46; DB 2; Length 1891;
Best Local Similarity 72.7%; Pred. No. 3e+02;
Matches 8; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 10 AGWGGGKRV 20
Db 1372 AGWGGGQSV 1382

RESULT 12

US-08-804-198-6
Sequence 6, Application US/08804198
Patent No. 5945320

GENERAL INFORMATION:

APPLICANT: Burgett, Stanley G.
APPLICANT: Kustoss, Stuart A.
APPLICANT: Rao, Nagaraja R.
APPLICANT: Richardson, Mark A.
APPLICANT: Rostock, Paul R., Jr.
TITLE OF INVENTION: PLATEINOLIDE SYNTHASE GENE
NUMBER OF SEQUENCES: 6
CORRESPONDENCE ADDRESS:
ADDRESSEE: PAUL R. CANTRELL 1138
STREET: LILLY CORPORATE CENTER
CITY: INDIANAPOLIS
STATE: IN

COUNTRY: USA

ZIP: 46285

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

COMPUTER: Macintosh

OPERATING SYSTEM: Macintosh 7.0

SOFTWARE: Microsoft Word 5.1

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/804,198

FILING DATE:

CLASSIFICATION: 435

ATTORNEY/AGENT INFORMATION:

NAME: CANTRELL, PAUL R.

REGISTRATION NUMBER: 36,470

REFERENCE/DOCKET NUMBER: P9113

TELECOMMUNICATION INFORMATION:

TELEPHONE: 317-276-3885

INFORMATION FOR SEQ ID NO: 6:

SEQUENCE CHARACTERISTICS:

LENGTH: 1891 amino acids

TYPE: amino acid

TOPOLOGY: unknown

MOLECULE TYPE: peptide

Query Match

Best Local Similarity 43.8%; Score 46; DB 2; Length 1891;
Best Local Similarity 72.7%; Pred. No. 3e+02;
Matches 8; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 10 AGWGGGKRV 20
Db 1372 AGWGGGQSV 1382

RESULT 13

US-09-428-517-4
Sequence 4, Application US/09428517
Patent No. 6251636

GENERAL INFORMATION:

APPLICANT: Betlach, Mary C.

APPLICANT: Shah, Sanjay Krishnakant

APPLICANT: McDaniel, Robert

APPLICANT: Tang, Li

TITLE OF INVENTION: RECOMBINANT OLEANDOLIDE POLYKETIDE SYNTHASE

FILE REFERENCE: 30062-20029.00

CURRENT APPLICATION NUMBER: US/09/428,517

CURRENT FILING DATE: 1999-10-28

EARLIER APPLICATION NUMBER: 60/120,254

EARLIER FILING DATE: 1999-02-16

EARLIER APPLICATION NUMBER: 60/106,100

EARLIER FILING DATE: 1998-10-29

NUMBER OF SEQ ID NOS: 12

SOFTWARE: Patentin Ver. 2.1

SEQ ID NO 4

LENGTH: 3519

TYPE: PRT

ORGANISM: Artificial Sequence

FEATURE:

OTHER INFORMATION: Description of Artificial Sequence: Recombinant

OTHER INFORMATION: Oleandolide PKs

US-09-428-517-4

Query Match

Best Local Similarity 43.8%; Score 46; DB 3; Length 3519;
Best Local Similarity 72.7%; Pred. No. 5.7e+02;
Matches 8; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 10 AGWGGGKRV 20
Db 1337 AGWGGGQSV 1347

RESULT 14

US-09-428-517-2

Sequence 2, Application US/09428517
Patent No. 6251636
GENERAL INFORMATION:
APPLICANT: Beilach, Mary C.
APPLICANT: Shah, Sanjay Krishnakant
APPLICANT: McDaniel, Robert
APPLICANT: Tang, Li
TITLE OF INVENTION: RECOMBINANT OLEANDOLIDE POLYKETIDE SYNTHASE
FILE REFERENCE: 30062-20029.00
CURRENT APPLICATION NUMBER: US/09/428,517
CURRENT FILING DATE: 1999-10-28
EARLIER APPLICATION NUMBER: 60/120,254
EARLIER FILING DATE: 1999-02-16
EARLIER APPLICATION NUMBER: 60/106,100
EARLIER FILING DATE: 1998-10-29
NUMBER OF SEQ ID NOS: 12
SOFTWARE: Patentin Ver. 2.1
SEQ ID NO 2
LENGTH: 4150
TYPE: PRT
ORGANISM: Artificial Sequence
FEATURE:
OTHER INFORMATION: Description of Artificial Sequence: Recombinant
OTHER INFORMATION: Oleandolide PKS
US-09-428-517-2

Query Match 43.8%; Score 46; DB 3; Length 4150;
Best Local Similarity 72.7%; Pred. No. 6.7e+02;
Matches 8; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 10 AGVWGGGKRV 20
Db 3838 AGVWGGGQAV 3848

RESULT 15
US-08-615-279-26
Sequence 26, Application US/08615279
Patent No. 5804371
GENERAL INFORMATION:
APPLICANT: H SS, Eva
APPLICANT: SEIDEL, Christoph
APPLICANT: WIENHUES, Ursula-Henrike
APPLICANT: PAATZ, Elke
APPLICANT: SCHMITT, Urban
TITLE OF INVENTION: HAPTEN-LABELLED PEPTIDES
NUMBER OF SEQUENCES: 40
CORRESPONDENCE ADDRESS:
ADDRESSEE: NIKAIKO, MARIELSTEIN, MURRAY & ORAM LLP
STREET: 655 Fifteenth Street, N. W., Suite 330 - G
CITY: Washington
STATE: DC
COUNTRY: USA
ZIP: 20005-5701
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/615,279
FILING DATE: 25-MAR-1996
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: PCT/EP95/02921
FILING DATE: 24-JUL-1995
ATTORNEY/AGENT INFORMATION:
NAME: BERNAN, Richard J.
REGISTRATION NUMBER: 39,107
REFERENCE/DOCKET NUMBER: P564-6006
TELECOMMUNICATION INFORMATION:
TELEPHONE: 202/638-5000

TELEFAX: 202/638-4810
INFORMATION FOR SEQ ID NO: 26:
SEQUENCE CHARACTERISTICS:
LENGTH: 32 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: protein
FEATURE:
NAME/KEY: Modified-site
LOCATION: 1
OTHER INFORMATION: /note=
OTHER INFORMATION: "digoxigenin-3-cme-beta-alanine-epsilon-aminocaproic
OTHER INFORMATION: acid-beta-alanine is attached to the Leu at the 1 position."
FEATURE:
NAME/KEY: Modified-site
LOCATION: 26
OTHER INFORMATION: /note="cme-3-digoxigenin is
OTHER INFORMATION: attached to the Leu at the 26 position."
US-08-615-279-26

Query Match 42.9%; Score 45; DB 1; Length 32;
Best Local Similarity 36.8%; Pred. No. 7;
Matches 7; Conservative 5; Mismatches 7; Indels 0; Gaps 0;

QY 2 ELVFTGYLAGVWGGGKRV 20
Db 9 KLICTFVLISWGCSSGLI 27

Search completed: September 30, 2004, 11:00:05
Job time: 13.9048 secs

This Page Blank (uspto)

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: September 30, 2004, 10:58:43 : Search time 42.0952 Seconds
(without alignments)
149.907 Million cell updates/sec

Title: US-09-988-851A-7
Perfect score: 105
Sequence: 1 IEHFTGVLAGVWGEGKRFV 20

Scoring table: BIOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 1017041 seqs, 315518202 residues

Total number of hits satisfying chosen parameters: 1017041

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%
Listing first 45 summaries

Database :
1: SP archaea:*
2: SP bacteria:*
3: SP fungi:*
4: SP human:*
5: SP invertebrate:*
6: SP mammal:*
7: SP msc:*
8: SP organelle:*
9: SP phage:*
10: SP plant:*
11: SP rodent:*
12: SP virus:*
13: SP vertebrate:*
14: SP unclassified:*
15: SP virus:*
16: SP bacteriopl:*
17: SP archaeap:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	60	57.1	761	12	0806C1
2	55	52.4	226	16	067094
3	55	50.5	294	16	07UEN7
4	52	49.5	255	16	025388
5	52	49.5	255	16	09ZUG1
6	52	49.5	270	17	08PT69
7	52	49.5	709	10	08IP19
8	52	49.5	721	10	08ZUT2
9	50	47.6	279	2	0934L5
10	50	47.6	328	16	09A2Y5
11	49.5	47.1	208	16	08XEM4
12	49	46.7	251	16	08XBB8
13	49	46.7	251	16	08NYV7
14	48.5	46.2	1485	2	084IT5
15	48	45.7	319	16	08DID3
16	48	45.7	3816	2	09KIV3

17	47	44.8	160	16	087HV3	Q87HV3 vibrio para
18	47	44.8	261	17	026968	Q26968 mechanobact
19	47	44.8	263	16	082LH6	Q82LH6 streptomyc
20	47	44.8	277	16	082R29	Q82R29 nitrosomona
21	47	44.8	343	2	Q81019	Q81019 gamma-prote
22	47	44.8	344	2	Q81052	Q81052 gamma-prote
23	47	44.8	344	2	Q810G4	Q810G4 gamma-prote
24	47	44.8	348	2	Q810J7	Q810J7 gamma-prote
25	47	44.8	348	2	Q8KZX1	Q8KZX1 gamma-prote
26	47	44.8	353	2	Q81048	Q81048 gamma-prote
27	47	44.8	353	2	Q81013	Q81013 gamma-prote
28	47	44.8	353	2	Q8KZY3	Q8KZY3 gamma-prote
29	47	44.8	353	2	Q81033	Q81033 gamma-prote
30	47	44.8	353	2	Q810D8	Q810D8 gamma-prote
31	47	44.8	353	2	Q81020	Q81020 gamma-prote
32	47	44.8	353	2	Q81015	Q81015 gamma-prote
33	47	44.8	353	2	Q810H6	Q810H6 gamma-prote
34	47	44.8	353	2	Q8KZM9	Q8KZM9 gamma-prote
35	47	44.8	353	2	Q810L8	Q810L8 gamma-prote
36	47	44.8	353	2	Q8KZV6	Q8KZV6 gamma-prote
37	47	44.8	353	2	Q8KZX2	Q8KZX2 gamma-prote
38	47	44.8	353	2	Q8KZX0	Q8KZX0 gamma-prote
39	47	44.8	353	2	Q81009	Q81009 gamma-prote
40	47	44.8	353	2	Q81057	Q81057 gamma-prote
41	47	44.8	353	2	Q8KZX5	Q8KZX5 gamma-prote
42	47	44.8	353	2	Q81076	Q81076 gamma-prote
43	47	44.8	353	2	Q810H0	Q810H0 gamma-prote
44	47	44.8	353	2	Q8KZX4	Q8KZX4 gamma-prote
45	47	44.8	353	2	Q8KZX4	Q8KZX4 gamma-prote

ALIGNMENTS

RESULT 1
ID 0806C1 PRELIMINARY; PRT; 761 AA.
AC 0806C1;
AT 01-JUN-2003 (TEMBLrel. 24, Created)
DT 01-JUN-2003 (TEMBLrel. 24, Last sequence update)
DT 01-OCT-2003 (TEMBLrel. 25, Last annotation update)
DE Component of DNA helicase-primase complex (Component of
DE helicase-primase complex).
GN US8
OS Simian herpes B virus (Cercopithecoid herpesvirus 1) (Shbv).
OC Viruses; dsDNA viruses, no RNA stage; Herpesviridae;
OC Alphaherpesvirinae; Simplexvirus.
CX NCBI_TaxID=10325;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=2267624; PubMed=12721804;
RA Ohsawa K., Black D.H., Sato H., Rogers K., Eberle R.;
RT "Sequence and genetic arrangement of the UL region of the monkey B
RT virus (Cercopithecine herpesvirus 1) genome and comparison with the UL
RT region of other primate herpesviruses.";
RL Arch. Virol. 148:989-997(2003).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=B2490;
RX MEDLINE=22628476; PubMed=12743273;
RA Pereleygina L., Zhu L., Zurkhlen H., Mills R., Borodovsky M.,
RA Hilliard J.K.;
RT "Complete Sequence and Comparative Analysis of the Genome of Herpes B
RT Virus (Cercopithecine Herpesvirus 1) from a Rhesus Monkey.";
RL J. Virol. 77:6167-6177(2003).
DR EMBL; AB096160; BAC58947.1; -;
DR EMBL; AF533768; AAF41426.1; -;
DR GO; GO:0004386; F:helicase activity; IEA.
DR GO; GO:0015079; F:viral genome replication; IEA.
DR InterPro; IPR004996; Herpes_HBPA.
DR Pfam; PF03324; Herpes_HBPA.1.
KW Helicase.
SQ SEQUENCE 761 AA; 79728 MW; 9AD9F9B66CF8915E CRC64;

Query Match 57.1%; Score 60; DB 12; Length 761;
 Best Local Similarity 57.9%; Pred. No. 2.9;
 Matches 11; Conservative 2; Mismatches 6; Indels 0; Gaps 0;

Qy 1 TELVFTGLAGWEGGK 19
 Db 720 MEAVFRGALTLGMEGCGF 738

RESULT 2

067094 PRELIMINARY; PRT; 226 AA.
 AC 067094;
 DT 01-AUG-1998 (TREMBLrel. 07, Created)
 DT 01-AUG-1998 (TREMBLrel. 07, Last sequence update)
 DT 01-JUN-2003 (TREMBLrel. 24, Last annotation update)
 DE Hypothetical protein A0_963.
 GN A0_963.
 OS Aquifex aeolicus.
 OC Bacteria; Aquificae; Aquificales; Aquificaceae; Aquifex.
 OX NCBI_TaxID=63363;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=VF5;
 RX MEDLINE=98196666; PubMed=9537320;
 RA Deckert G., Warren P.V., Gaasterland T., Young W.G., Lenox A.L.,
 RA Graham D.E., Overbeek R., Sneed M.A., Keller M., Aubay M., Huber R.,
 RA Feldman R.A., Short J.M., Olson G.J., Swanson R.V.;
 RT "The complete genome of the hyperthermophilic bacterium Aquifex
 aeolicus";
 RL Nature 392:353-358 (1998).
 DR EMBL; AE000715; AAC07057.1; -.
 DR PIR; C70383; C70383.
 KW Hypothetical protein; Complete proteome.
 SQ SEQUENCE 226 AA; 25939 MW; 29D45D27634281B CRC64;

Query Match 52.4%; Score 55; DB 16; Length 226;
 Best Local Similarity 50.0%; Pred. No. 4.5;
 Matches 9; Conservative 4; Mismatches 5; Indels 0; Gaps 0;

Qy 3 LVFTGLAGWEGGK 20
 Db 163 VFTGLANMAGGNY 180

RESULT 3

07UEN7 PRELIMINARY; PRT; 294 AA.
 AC 07UEN7;
 DT 01-OCT-2003 (TREMBLrel. 25, Created)
 DT 01-OCT-2003 (TREMBLrel. 25, Last sequence update)
 DT 01-OCT-2003 (TREMBLrel. 25, Last annotation update)
 DE Hypothetical protein.
 GN RB11226.
 OS Rhodospirillum rubrum.
 OC Bacteria; Planctomycetes; Planctomycetacia; Planctomycetales;
 OC Planctomycetaceae; Pirellula.
 OX NCBI_TaxID=117;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=1;
 RX MEDLINE=2735913; PubMed=12835416;
 RA Gleckner F.O., Kube M., Bauer M., Teeling H., Lombardot T.,
 RA Ludwig W., Gade D., Beck A., Borzym K., Heilmann K., Rabus R.,
 RA Schleutner H., Amann R., Reinhardt R.;
 RT "Complete genome sequence of the marine planctomycete Pirellula sp.
 strain 1";
 RL Proc. Natl. Acad. Sci. U.S.A. 100:8298-8303 (2003).
 DR EMBL; BX294153; CAD78998.1; -.
 KW Hypothetical protein; Complete proteome.
 SQ SEQUENCE 294 AA; 30660 MW; 55C8B8A1A78F8790 CRC64;

Query Match 50.5%; Score 53; DB 16; Length 294;
 Best Local Similarity 58.8%; Pred. No. 12;
 Matches 10; Conservative 4; Mismatches 3; Indels 0; Gaps 0;

Qy 4 VFTGLAGWEGGK 20
 Db 183 LMTGLSLFVGGGF 199

RESULT 4

025388 PRELIMINARY; PRT; 255 AA.
 AC 025388;
 DT 01-JAN-1998 (TREMBLrel. 05, Created)
 DT 01-JAN-1998 (TREMBLrel. 05, Last sequence update)
 DT 01-JUN-2003 (TREMBLrel. 24, Last annotation update)
 DE Conserved hypothetical integral membrane protein.
 GN HP067.
 OS Helicobacter pylori (Campylobacter pylori).
 OC Bacteria; Proteobacteria; Epsilonproteobacteria; Campylobacteriales;
 OC Helicobacteraceae; Helicobacter.
 OX NCBI_TaxID=210;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=26595 / ATCC 700392;
 RX MEDLINE=97394467; PubMed=9252185;
 RA Tomb J.-F., White O., Kerlavage A.R., Clayton R.A., Sutton G.G.,
 RA Fleischmann R.D., Ketchum K.A., Klein H.-P., Gill S., Dougherty B.A.,
 RA Nelson K., Quackenbush J., Zhou L., Kirness E.F., Peterson S.,
 RA Loftus B., Richardson D., Dodson R., Khairat H.G., Glodek A.,
 RA McKenney K., Fitzgerald L.M., Lee N., Adams M.D., Hickey E.K.,
 RA Berg D.E., Gocayne J.D., Utterback T.R., Peterson J.P., Kelley J.M.,
 RA Cotton M.D., Melman J.N., Fujii C., Bowman C., Watney L., Wallin E.,
 RA Hayes W.S., Borodovsky M., Katp P.D., Smith H.O., Fraser C.M.,
 RA Venter J.C.;
 RT "The complete genome sequence of the gastric pathogen Helicobacter
 pylori";
 RL Nature 388:539-547 (1997).
 DR EMBL; AE000581; AAD14882.1; -.
 DR PIR; E64604; E64604.
 DR TIGR; HP0677; -.
 DR InterPro; IPR002781; DUF81.
 DR InterPro; IPR001092; HLH_basic.
 DR Pfam; PF01925; DUF81.1.
 DR PROSITE; PS00038; HLH_1; 1.
 KW Hypothetical protein; Complete proteome.
 SQ SEQUENCE 255 AA; 27432 MW; 669C9968230ABE7E CRC64;

Query Match 49.5%; Score 52; DB 16; Length 255;
 Best Local Similarity 47.1%; Pred. No. 14;
 Matches 8; Conservative 6; Mismatches 3; Indels 0; Gaps 0;

Qy 4 VFTGLAGWEGGK 20
 Db 12 LFTGLSLFVGGGF 28

RESULT 5

09ZLG1 PRELIMINARY; PRT; 255 AA.
 AC 09ZLG1;
 DT 01-MAY-1999 (TREMBLrel. 10, Created)
 DT 01-MAY-1999 (TREMBLrel. 10, Last sequence update)
 DT 01-JUN-2003 (TREMBLrel. 24, Last annotation update)
 DE Putative.
 GN JHP0619.
 OS Helicobacter pylori J99 (Campylobacter pylori J99).
 OC Bacteria; Proteobacteria; Epsilonproteobacteria; Campylobacteriales;
 OC Helicobacteraceae; Helicobacter.
 OX NCBI_TaxID=85963;
 RN [1]
 RP SEQUENCE FROM N.A.
 SQ SEQUENCE 255 AA; 27432 MW; 669C9968230ABE7E CRC64;

```

RA Alm R.A., Ling L.-S.L., Moir D.T., King B.L., Brown E.D., Doig P.C.,
RA Smith D.R., Noonan B., Gulid B.C., deJonge B.L., Carmel G.,
RA Thumming P.J., Caruso A., Uria-Nickelsen M., Mills D.M., Ives C.,
RA Gibson R., Merberg D., Mills S.D., Jiang Q., Taylor D.E., Vovis G.F.,
RA Trust T.J.;
RT "Genomic sequence comparison of two unrelated isolates of the human
RT gastric pathogen Helicobacter pylori.";
RL Nature 397:176-180(1999).
DR EMBL: AB001493; AAD06192.1; -.
DR F.R. D71910; D71910.
DR InterPro: IPR002781; DUF81.
DR InterPro: IPR001092; HLH_basic.
DR Pfam: PF01925; DUF81; 1.
DR PROSITE: PS00038; HMH_1; 1.
KW Complete proteome.
SQ SEQUENCE 255 AA; 27492 MW; D85430305BD7847 CRC64;

Query Match 49.5%; Score 52; DB 16; Length 255;
Best Local Similarity 47.1%; Pred. No. 14;
Matches 8; Conservative 6; Mismatches 3; Indels 0; Gaps 0;

QY 4 FTGVLGAGWGEQGV 20
Db 12 FTGILSGIFGIGGMI 28

RESULT 6
ID Q8PT69 PRELIMINARY; PRT; 270 AA.
AC Q8PT69;
DT 01-OCT-2002 (TRENBLrel. 22, Created)
DT 01-OCT-2002 (TRENBLrel. 22, Last sequence update)
DT 01-OCT-2002 (TRENBLrel. 22, Last annotation update)
DE Hypothetical permealase protein MM2849.
GN MM2849;
OS Archaea; Euryarchaeota; Euryarchaeota orders incertae sedis;
OC Methanosarcinales; Methanosarcinaceae; Methanosarcina.
OC NCBI_TaxID=2209;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=Geel / G01 / ATCC BAA-199 / DSM 3647 / OCM 88;
RX DEPMELINE=22120827; PubMed=12125824;
RA Depmelter U., Johann A., Hartisch T., Werkl R., Schmitz R.A.,
RA Martinez-Artis R., Henne A., Wierzer A., Baeumer S., Jacobi C.,
RA Brueggemann H., Lienard T., Christmann A., Boencke M., Steckel S.,
RA Bhattacharya A., Lykdis A., Overbeek R., Klenk H.-P., Gunsalus R.P.,
RA Fritz H.-J., Gottschalk G.;
RT "The genome of Methanosarcina mazei: evidence for lateral gene
RT transfer between Bacteria and Archaea.";
RL J. Mol. Microbiol. Biotechnol. 4:453-461(2002).
DR EMBL: AE013535; AAM32545.1; -.
DR InterPro: IPR002781; DUF81.
DR Pfam: PF01925; DUF81; 1.
KW Complete proteome; Hypothetical protein.
SQ SEQUENCE 270 AA; 28911 MW; 44F1BD612B564A82 CRC64;

Query Match 49.5%; Score 52; DB 17; Length 270;
Best Local Similarity 56.2%; Pred. No. 15;
Matches 9; Conservative 4; Mismatches 3; Indels 0; Gaps 0;

QY 5 FTGVLGAGWGEQGV 20
Db 18 FTGVLGAGWGEQGV 33

RESULT 7
ID Q8PLP9 PRELIMINARY; PRT; 709 AA.
AC Q8PLP9;
DT 01-OCT-2002 (TRENBLrel. 22, Created)
DT 01-OCT-2002 (TRENBLrel. 22, Last sequence update)
DT 01-OCT-2002 (TRENBLrel. 22, Last annotation update)
DT 01-JUN-2003 (TRENBLrel. 24, Last annotation update)

```

```

DE Putative membrane transporter.
GN AT2G27810.
OS Arabidopsis thaliana (Mouse-ear cress).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids;
OC eurosid II; Brassicales; Brassicaceae; Arabidopsids.
OX NCBI_TaxID=3702;
RN [1]
RP SEQUENCE FROM N.A.
RA Nguyen M., Karlin-Neumann G., Southwick A., Lam B., Miranda M.,
RA Palm C.J., Bowser L., Jones T., Banh U., Caminici P., Chen H.,
RA Cheuk R., Chung M.K., Hayashizaki Y., Ishida J., Kamiya A., Kawai J.,
RA Kim C., Lin J., Liu S.X., Narusaka M., Pham P.K., Sakano H.,
RA Sakurai T., Satou M., Seki M., Shinn P., Yamada K., Shinozaki K.,
RA Ecker J., Theologis A., Davis R.W.;
RL Submitted (APR-2002) to the EMBL/GenBank/DBJ databases.
RN [2]
RP SEQUENCE FROM N.A.
RA Nguyen M., Karlin-Neumann G., Southwick A., Tripp M., Miranda M.,
RA Palm C.J., Bowser L., Jones T., Banh U., Caminici P., Chen H.,
RA Cheuk R., Chung M.K., Hayashizaki Y., Ishida J., Kamiya A., Kawai J.,
RA Kim C., Lin J., Liu S.X., Narusaka M., Pham P.K., Sakano H.,
RA Sakurai T., Satou M., Seki M., Shinn P., Yamada K., Shinozaki K.,
RA Ecker J., Theologis A., Davis R.W.;
RL Submitted (NOV-2002) to the EMBL/GenBank/DBJ databases.
DR EMBL: AY095945; AAM20397.1; -.
DR EMBL: BT002121; AAM72132.1; -.
DR GO: GO:0005622; C:intracellular; IEA.
DR GO: GO:0016020; C:membrane; IEA.
DR GO: GO:0003700; F:transcription factor activity; IEA.
DR GO: GO:0005215; F:transporter activity; IEA.
DR GO: GO:0006355; P:regulation of transcription, DNA-dependent; IEA.
DR GO: GO:0006810; P:transport; IEA.
DR InterPro: IPR000005; HTHRAC.
DR Pfam: PF00660; xan_ux_permease; 1.
DR PROSITE: PS00041; HTX_ARAC_FAMILY_1; 1.
SQ SEQUENCE 709 AA; 76674 MW; BD6CE057C49EF3FB CRC64;

Query Match 49.5%; Score 52; DB 10; Length 709;
Best Local Similarity 75.0%; Pred. No. 40;
Matches 9; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 5 FTGVLGAGWGEQ 16
Db 492 FTGVLGAGWGTG 503

RESULT 8
ID Q9ZUY2 PRELIMINARY; PRT; 721 AA.
AC Q9ZUY2;
DT 01-MAY-1999 (TRENBLrel. 10, Created)
DT 01-MAY-1999 (TRENBLrel. 10, Last sequence update)
DT 01-JUN-2003 (TRENBLrel. 24, Last annotation update)
DE Putative membrane transporter.
GN AT2G27810.
OS Arabidopsis thaliana (Mouse-ear cress).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids;
OC eurosid II; Brassicales; Brassicaceae; Arabidopsids.
OX NCBI_TaxID=3702;
RN [1]
RP SEQUENCE FROM N.A.
RA STRAIN=cv. Columbia;
RX MEDLINE=20083487; PubMed=16617197;
RA Lin X., Kaul S., Rounsley S.D., Shea T.P., Benito M.-I., Town C.D.,
RA Fujii C.Y., Mason T.M., Bowman C.L., Barnstead M.E., Feldblum T.V.,
RA Buell C.R., Ketchum K.A., Lee J.J., Ronning C.M., Koo H., Moffat K.S.,
RA Cronin L.A., Shen M., Vanden S.E., Unagay L., Tallon L.J., Gill J.E.,
RA Adams M.D., Carrara A.J., Creasy T.H., Goodman H.M., Semeriville C.R.,
RA Copenhagen G.P., Preuss D., Nierman W.C., White O., Eisen J.A.,
RA Salzberg S.L., Fraser C.M., Venter J.C.;

```

```

RT "Sequence and analysis of chromosome 2 of the plant Arabidopsis
RL thaliana."
RL Nature 402:761-768 (1999).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=cv. Columbia;
RA Lin X.;
RL Submitted (MAR-2000) to the EMBL/GenBank/DBJ databases.
DR EMBL; AC005824; AAC73019.1; -.
DR PIR; C84677; C84677.
DR GO; GO:0005622; C:intracellular; IEA.
DR GO; GO:0016020; C:membrane; IEA.
DR GO; GO:0003700; F:transcription factor activity; IEA.
DR GO; GO:0005215; F:transporter activity; IEA.
DR GO; GO:0006355; P:regulation of transcription, DNA-dependent; IEA.
DR GO; GO:0006810; P:transport; IEA.
DR InterPro; IPR000005; HTHA3C.
DR InterPro; IPR006043; Xant/urac/vitc.
DR Pfam; PF00860; xan_ur permease; 1.
DR PROSITE; PS00041; HTH_ARAC_FAMILY_1; 1.
SQ SEQUENCE 721 AA; 78157 MW; 2317AE4FC3A9438 CRC64;

Query Match 49.5%; Score 52; DB 10; Length 721;
Best Local Similarity 75.0%; Pred. No. 41;
Matches 9; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

CY 5 FTGVLAVGWGEG 16
DB 492 FTSLAGLWGTG 503

RESULT 9
ID Q934L5 PRELIMINARY; PRT; 279 AA.
AC Q934L5;
DT 01-DEC-2001 (TREMBlrel. 19, Created)
DT 01-DEC-2001 (TREMBlrel. 19, Last sequence update)
DT 01-JUN-2003 (TREMBlrel. 24, Last annotation update)
DE Hypothetical integral membrane protein.
OS Vibrio salmonicida.
OC Bacteria; Proteobacteria; Gammaproteobacteria; Vibrionales;
OC Vibrionaceae; Vibrio.
OX NCBI_TaxID=40269;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=VS224;
RA Sorum H., Dommasnes K., Fjeldahl I., Berg I., Alvheim K.;
RT "A novel composite transposon in Vibrio salmonicida contains
RT methionine biosynthetic genes and an integron."
RL Submitted (MAY-2000) to the EMBL/GenBank/DBJ databases.
DR EMBL; AJ278113; CAC81936.1; -.
DR GO; GO:0046821; C:extrachromosomal DNA; IEA.
DR InterPro; IPR002781; DUF81.
DR Pfam; PF01925; DUF81; 1.
KM Plasmid.
SQ SEQUENCE 279 AA; 28766 MW; CC5136515601BD4 CRC64;

Query Match 47.6%; Score 50; DB 2; Length 279;
Best Local Similarity 58.8%; Pred. No. 30;
Matches 10; Conservative 2; Mismatches 5; Indels 0; Gaps 0;

CY 4 VFTGVLAVGWGEGKFV 20
DB 22 VFAGTLAVGLVGGIV 38

RESULT 10
ID Q9A2Y5 PRELIMINARY; PRT; 328 AA.
AC Q9A2Y5;
DT 01-JUN-2001 (TREMBlrel. 17, Created)
DT 01-JUN-2001 (TREMBlrel. 17, Last sequence update)

```

```

DT 01-JUN-2003 (TREMBlrel. 24, Last annotation update)
DE 4-hydroxybenzoate octaprenyltransferase.
GN CC3421.
OS Caulobacter crescentus.
OC Bacteria; Proteobacteria; Alphaproteobacteria; Caulobacteriales;
OC Caulobacteriaceae; Caulobacter.
OX NCBI_TaxID=155892;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=ATCC 19089 / CB15;
DR MEDLINE=21173698; PubMed=11259647;
RA Nierman W.C., Feldblum T.V., Laub M.T., Paulsen I.T., Nelson K.E.,
RA Eisen J., Heidelberg J.F., Alley M.R.K., Ohta N., Madcock J.R.,
RA Potocka I., Nelson W.C., Newton A., Stephens C., Phadke N.D., Ely B.,
RA DeBoy R.T., Dodson R.J., Durkin A.S., Gwinn M.L., Haft D.H.,
RA Kolony J.F., Smit J., Craven M.B., Khouli H., Shetty J., Berry K.,
RA Usterback T., Tran K., Wolf A., Vamathevan J., Ermolaeva M., White O.,
RA Salzberg S.L., Venter J.C., Shapiro L., Fraser C.M.;
RT "Complete genome sequence of Caulobacter crescentus."
RL Proc. Natl. Acad. Sci. U.S.A. 98:4136-4141 (2001).
DR EMBL; AE006002; AAK25383.1; -.
DR PIR; C87673; C87673.
DR TIGR; CC3421; -.
DR GO; GO:0016021; C:integral to membrane; IEA.
DR GO; GO:0004659; F:prenyltransferase activity; IEA.
DR GO; GO:0016740; F:transferase activity; IEA.
DR GO; GO:0009058; F:biosynthesis; IEA.
DR InterPro; IPR000537; Ubia.
DR InterPro; IPR006370; Ubia_proteo.
DR Pfam; PF01040; Ubia; 1.
DR TIGRPFAM; TIGR01474; ubia_proteo; 1.
DR PROSITE; PS00343; Ubia; 1.
DR Transferase; Complete proteome.
SQ SEQUENCE 328 AA; 35555 MW; 5D2671B8BEC32FA8 CRC64;

Query Match 47.6%; Score 50; DB 16; Length 328;
Best Local Similarity 52.9%; Pred. No. 36;
Matches 9; Conservative 3; Mismatches 5; Indels 0; Gaps 0;

CY 3 LVFTGVLAVGWGEGKF 19
DB 312 LVFAGVLAVGLWPGVSF 328

RESULT 11
ID Q8EXE4 PRELIMINARY; PRT; 208 AA.
AC Q8EXE4;
DT 01-MAR-2003 (TREMBlrel. 23, Created)
DT 01-MAR-2003 (TREMBlrel. 23, Last sequence update)
DT 01-MAR-2003 (TREMBlrel. 23, Last annotation update)
DE Conserved hypothetical protein.
GN LB269.
OS Leptospira interrogans.
OC Bacteria; Spirochaetes; Spirochaetales; Leptospiraceae; Leptospira.
OX NCBI_TaxID=1173;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=56601 / Serogroup Icterohaemorrhagiae / Serovar lai;
RA Ren S.;
RL Submitted (MAR-2002) to the EMBL/GenBank/DBJ databases.
DR EMBL; AE011614; AAN51828.1; -.
KM Hypothetical protein; Complete proteome.
SQ SEQUENCE 208 AA; 23084 MW; 458BFFA9FFFA13A CRC64;

Query Match 47.1%; Score 49.5; DB 16; Length 208;
Best Local Similarity 52.6%; Pred. No. 26;
Matches 10; Conservative 3; Mismatches 5; Indels 1; Gaps 1;

CY 3 LVFTGVLAVGW-GEQKRFV 20
DB 38 LVFFSFLGKWKKGKGRFI 56

```


RESULT 12

Q9XB8 PRELIMINARY; PRT; 251 AA.
 ID Q9XB8;
 AC Q9XB8;
 DT 01-NOV-1999 (TREMblrel. 12, Created)
 DT 01-NOV-1999 (TREMblrel. 12, Last sequence update)
 DT 01-JUN-2003 (TREMblrel. 24, Last annotation update)
 DE ORF N002 (Hypothetical protein SA0080) (Hypothetical protein SA0080)
 GN SAV0084 OR SA0080.
 OS Staphylococcus aureus (strain Mu50 / ATCC 700699),
 OS Staphylococcus aureus (strain N315), and
 OS Staphylococcus aureus.
 OC Bacteria; Firmicutes; Bacillales; Staphylococcus.
 NX NCBI_TaxID=158878, 158879, 1280;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC SPECIES=S.aureus; TRANSPON=TN55X;
 RX MEDLINE=99278010; PubMed=10348769;
 RA Ito T., Katayama Y., Hiramatsu K.;
 RT "Cloning and nucleotide sequence determination of the entire mec DNA
 of pre-mechicillin-resistant Staphylococcus aureus N315.";
 RL Antimicrob. Agents Chemother. 43:1449-1458(1999).
 RN [2]
 RP SEQUENCE FROM N.A.
 RC SPECIES=S.aureus; STRAIN=NCTC10442;
 RA Ito T., Hiramatsu K.;
 RT "Comparison of three types of SRmec.";
 RL Submitted (Oct-1999) to the EMBL/Genbank/DBJ databases.
 RN [3]
 RP SEQUENCE FROM N.A.
 RC SPECIES=S.aureus (strain Mu50), and S.aureus (strain N315);
 RX MEDLINE=21311952; PubMed=11418146;
 RA Kuroda M., Ohts T., Uchiyama I., Baba T., Yuzawa H., Kobayashi I.,
 RA Cui L., Oguchi A., Aoki K.-I., Nagai Y., Iian J.-Q., Ito T.,
 RA Katamori M., Matsumaru H., Matuyama A., Murakami H., Hosoyama A.,
 RA Mizutani-U Y., Takahashi N.K., Sawano T., Inoue R.-I., Kaito C.,
 RA Sekimizu K., Hirakawa H., Kuhara S., Goto S., Yabuzaki U.,
 RA Kanehisa M., Yamashita A., Oshima K., Furuya K., Yoshino C., Shiba T.,
 RA Hattori M., Ogasawara N., Hayashi H., Hiramatsu K.;
 RT "Whole genome sequencing of methicillin-resistant Staphylococcus
 aureus.";
 RL Lancet 357:1225-1240(2001).
 DR EMBL; D86934; BA8212.1; -;
 DR EMBL; AB033763; BA8663.1; -;
 DR EMBL; AP003358; BAB56246.1; -;
 DR EMBL; AP003129; BAB41299.1; -;
 DR PIR; T44070; T44070.
 DR InterPro; IPR002781; DUF81.
 DR Pfam; PF01925; DUF81.1.
 KM Hypothetical protein; Complete proteome.
 SQ SEQUENCE 251 AA; 2690 MW; E4D47733BFL668A3 CRC64;

Query Match 46.7%; Score 49; DB 16; Length 251;
 Best Local Similarity 41.2%; Pred. No. 38;
 Matches 7; Conservative 6; Mismatches 4; Indels 0; Gaps 0;

QY 4 VFTGLAGVWGEGKFEV 20
 Db 146 LFIGTISGIVGAGAFI 162

OC Bacteria; Firmicutes; Bacillales; Staphylococcus.
 NX NCBI_TaxID=196620;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=22040717; PubMed=12044378;
 RA Baba T., Takeuchi F., Kuroda M., Yuzawa H., Aoki K.-I., Oguchi A.,
 RA Nagai Y., Iwama N., Asano K., Naim T., Kuroda H., Cui L.,
 RA Yamamoto K., Hiramatsu K.;
 RT "Genome and virulence determinants of high virulence community-
 RT acquired MRSA.";
 RL Lancet 359:1819-1827(2002).
 DR EMBL; AP004622; BAB93919.1; -;
 DR InterPro; IPR002781; DUF81.
 DR Pfam; PF01925; DUF81.1.
 KM Hypothetical protein; Complete proteome.
 SQ SEQUENCE 251 AA; 26704 MW; E4D47733AD8668A3 CRC64;

Query Match 46.7%; Score 49; DB 16; Length 251;
 Best Local Similarity 41.2%; Pred. No. 38;
 Matches 7; Conservative 6; Mismatches 4; Indels 0; Gaps 0;

QY 4 VFTGLAGVWGEGKFEV 20
 Db 146 LFIGTISGIVGAGAFI 162

RESULT 14
 Q84IT5 PRELIMINARY; PRT; 1485 AA.
 ID Q84IT5;
 AC Q84IT5;
 DT 01-JUN-2003 (TREMblrel. 24, Created)
 DT 01-JUN-2003 (TREMblrel. 24, Last sequence update)
 DT 01-OCT-2003 (TREMblrel. 25, Last annotation update)
 DE Salinomycin polyketide synthase (Fragment).
 GN SALA.
 OS Streptomyces albus.
 OC Bacteria; Actinobacteria; Actinobacteriales; Actinomycetales;
 OC Streptomycetaceae; Streptomycetaceae; Streptomycetes.
 NX NCBI_TaxID=1888;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=ATCC21838;
 RA Isumikawa M., Murata M., Tachibana K., Ebizuka Y., Fujii I.;
 RT "Cloning of Polyketide Synthase Genes Involved in Salinomycin
 RT Biosynthesis from Streptomyces albus.";
 RL Submitted (Jul-2002) to the EMBL/Genbank/DBJ databases.
 DR EMBL; AB087998; BAC54914.1; -;
 DR GO; GO:0016740; F:transferase activity; IEA.
 DR GO; GO:0005215; F:transporter activity; IEA.
 DR GO; GO:0006633; P:fatty acid biosynthesis; IEA.
 DR GO; GO:0008152; P:metabolism; IEA.
 DR GO; GO:0006810; P:transport; IEA.
 DR InterPro; IPR001227; AcTrans.
 DR InterPro; IPR000794; ketoacyl_synth.
 DR InterPro; IPR000566; lipopol_cycFABP.
 DR InterPro; IPR006162; Pantane_S.
 DR InterPro; IPR006163; P_bind.
 DR Pfam; PF00698; Acyl_transf. 1.
 DR Pfam; PF00109; ketoacyl_synth. 1.
 DR Pfam; PF02801; ketoacyl_synth_C. 1.
 DR Pfam; PF00550; PP-binding 1.
 DR PROSITE; PSS0075; ACP_DOMAIN; 1.
 DR PROSITE; PS00606; B_KETOACYL_SYNTHASE; 1.
 DR PROSITE; PS00213; LIPOCALIN; 1.
 DR PROSITE; PS00012; PHOSPHOPANTETHEINE; 1.
 FT NON_TER 1
 FT NON_TER 1
 SQ SEQUENCE 1485 AA; 15353 MW; 34AD09A885311BD CRC64;

Query Match 46.2%; Score 48.5; DB 2; Length 1485;
 Best Local Similarity 56.2%; Pred. No. 2.8e+02;
 Matches 9; Conservative 4; Mismatches 2; Indels 1; Gaps 1;

QY 3 LVFTGVLAGVWGEGGK 18
 :||:|||||:
 DB 1416 VFESSI-AGVWGSGGQ 1430

RESULT 15

Q8DIH3 PRELIMINARY; PRT; 319 AA.
 ID Q8DIH3
 AC Q8DIH3
 DT 01-MAR-2003 (TReMBLrel. 23, Created)
 DT 01-MAR-2003 (TReMBLrel. 23, Last sequence update)
 DT 01-JUN-2003 (TReMBLrel. 24, Last annotation update)
 DE C-type cytochrome synthesis protein.
 GN CCSA OR TLR1615.
 OS Synechococcus elongatus (Thermosynechococcus elongatus).
 OC Bacteria; Cyanobacteria; Chroococcales; Synechococcus.
 OX NCBI_TaxID=32046;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=BP-1;
 RX MEDLINE=2225144; PubMed=12240834;
 RA Nakamura Y., Kaneo T., Sato S., Ikeuchi M., Katoh H., Sasamoto S.,
 RA Watanabe A., Iriuchi M., Kawashima K., Kimura T., Kishida Y.,
 RA Kiyokawa C., Kohara M., Matsumoto M., Matsuno A., Nakazaki N.,
 RA Shimpō S., Sugimoto M., Takeuchi C., Yamada M., Tabata S.,
 RT "Complete genome structure of the thermophilic cyanobacterium
 Thermosynechococcus elongatus BP-1.";
 RL DNA Res. 9:123-130(2002).
 DR EMBL; AP005374; BAC09167.1; -.
 DR GO; GO:0016020; C:membrane; IEA.
 DR GO; GO:0008535; P:cytochrome c oxidase biogenesis; IEA.
 DR GO; GO:0006461; P:protein complex assembly; IEA.
 DR InterPro; IPR002541; CytC_asm.
 DR Pfam; PF01578; CytC_asm; 1.
 KW Complete proteome.
 SQ SEQUENCE 319 AA; 34633 MW; 69D334382587731D CRC64;
 Query March 45.7%; Score 48; DB 16; Length 319;
 Best Local Similarity 58.8%; Pred. No. 68;
 Matches 10; Conservative 1; Mismatches 6; Indels 0; Gaps 0;

QY 3 LVFTGVLAGVWGEGGK 19
 :||:|||||:
 DB 50 LCITGLAARWIEGGYF 66

Search completed: September 30, 2004, 11:08:34
 Job time : 44.0952 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: September 30, 2004, 10:56:43 ; Search time 7.42857 Seconds
(without alignments)
140.189 Million cell updates/sec

Title: US-09-988-851a-7

Perfect score: 105
Sequence: 1 IEIVFTGVLAGVWGEGKRV 20

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 141681 seqs, 52070155 residues

Total number of hits satisfying chosen parameters: 141681

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : SwissProt_42:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	105	100.0	750	1	HEPA_HSV1
2	92	87.6	752	1	HEPA_HSV1H
3	49	46.7	160	1	YCDZ_SALTY
4	46	43.8	163	1	YCDZ_ECOLI
5	46	43.8	417	1	CLCB_SALTY
6	45	43.8	3519	1	OL56_STRTAT
7	45	42.9	861	1	ENV_HV1K
8	44.5	42.4	536	1	DTI_YEAST
9	44	41.9	255	1	PABG_MYCMA
10	44	41.9	267	1	V441_METVA
11	44	41.9	444	1	GCSA_CHLVE
12	44	41.9	1520	1	PMPD_CHLNU
13	44	41.9	1901	1	YZ08_MYCTU
14	43.5	41.4	1882	1	ENV_STVM1
15	43	41.0	355	1	YAO0_PYRO
16	43	41.0	431	1	KDPA_CHLVR
17	43	41.0	628	1	PRPE_ECOLI
18	43	41.0	726	1	HEPA_THEMA
19	42.5	40.5	261	1	ZNTB_ECOLI
20	42.5	40.5	452	1	NIAM_BRAFL
21	42.5	40.5	452	1	NIAM_BRAFL
22	42	40.0	213	1	YAL7_HABIN
23	42	40.0	253	1	TRID_DROER
24	42	40.0	286	1	AL4_ASPFU
25	42	40.0	340	1	YV90_CLOAB
26	42	40.0	664	1	EMAP_LYTV
27	42	40.0	670	1	ACSA_PYRAE
28	42	40.0	666	1	EMAP_STRPU
29	42	40.0	1531	1	PMPD_CHLVR
30	41.5	39.5	358	1	DDJB_CLOTB
31	41.5	39.5	358	1	GATB_PHILO
32	41.5	39.5	500	1	GATB_BRUME
33	41.5	39.5	572	1	YBNI_SCHPO

ALIGNMENTS

34	41.5	39.5	573	1	YBNI_SCHPO	Q9p6j9 schizosacch
35	41	39.0	247	1	PABG_MYCTU	Q48930 mycobacteri
36	41	39.0	255	1	PABG_MYCAV	007399 mycobacteri
37	41	39.0	293	1	HRPX_YERPE	084fj7 yersinia pe
38	41	39.0	352	1	PURS_COXBU	Q83a20 coxiella bu
39	41	39.0	366	1	MWAY_CORGL	Q8nmn2 coynebacte
40	41	39.0	476	1	PURA_WHEAT	Q24396 triticum ae
41	41	39.0	484	1	PURA_MAIZE	Q24578 zea mays (m
42	41	39.0	511	1	SYR6_MOUSE	Q9j0c8 mus musculu
43	41	39.0	511	1	SVT6_RAT	Q62746 rattus norv
44	41	39.0	538	1	KPRI_RHIME	Q52938 rhizobium m
45	41	39.0	637	1	VATI_THERAC	Q9nm61 thermoplasm

RESULT 1
HEPA_HSV1
ID HEPA_HSV1 STANDARD; ERT; 750 AA.

AC P10192;
DT 01-MAR-1989 (Rel. 10, Created)
DT 01-MAR-1989 (Rel. 10, Last sequence update)
DT 01-OCT-1996 (Rel. 34, Last annotation update)
DE DNA helicase/primase complex associated protein.
GN UL8.
OS Herpes simplex virus (type 1 / strain 17).
OC Viruses; dsDNA viruses, no RNA stage; Herpesviridae;
OC Alphaherpesvirinae; Simplexvirus.
OX NCBI_TaxID=10239;
RN [1]
RX MEDLINE=88274327; PubMed=2839594;
RA McGeoch D.J., Dalrymple M.A., Davison A.J., Dolan A., Frame M.C.,
RA McNab D., Perry L.J., Scott J.E., Taylor P.,
RT "The complete DNA sequence of the long unique region in the genome of
RT herpes simplex virus type 1".
RL J. Gen. Virol. 69:1531-1574(1988).
RN [2]
RP SEQUENCE FROM N.A.
RX MEDLINE=88091053; PubMed=2826807;
RA McGeoch D.J., Dalrymple M.A., Dolan A., McNab D., Perry L.J.,
RA Taylor P., Challberg M.D.;
RT "Structures of herpes simplex virus type 1 genes required for
RT replication of virus DNA".
RL J. Virol. 62:444-453(1988).
CC -1- FUNCTION: PROBABLY INVOLVED IN DNA REPLICATION.
CC -1- SIMILARITY: BELONGS TO FAMILY THAT GROUPS TOGETHER HSV-1 UL8,
CC EBV-1 54, VZV 52 AND HCMV 102.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@sib-sib.ch).
CC
CC EMBL: X14112; CA33344.1; -;
CC EMBL: M19120; AAA45823.1; -;
CC PIR: C298900; WMBEX8.
CC InterPro: IPR004396; Herpes_HSPA.
CC Pfam: PF03324; Herpes_HSPA.1.
CC DNA replication.
CC SEQUENCE 750 AA; 79925 MW; EC3A9ABD085CB392 CRC64;

Query Match 100.0%; Score 105; DB 1; Length 750;
Best Local Similarity 100.0%; Pred. No. 3.6e-08;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 IEIVFTGVLAGVWGEGKRV 20
DB 719 IEIVFTGVLAGVWGEGKRV 738

RESULT 2
HEPA_HSV2H STANDARD; PRT; 752 AA.
ID HEPA_HSV2H
AC P8431;
DT 01-NOV-1997 (Rel. 35, Created)
DT 01-NOV-1997 (Rel. 35, Last sequence update)
DT 01-NOV-1997 (Rel. 35, Last annotation update)
DE DNA helicase/primase complex associated protein.
GN UL8
OS Herpes simplex virus (type 2 / strain H652).
OC Viruses; dsDNA viruses, no RNA stage; Herpesviridae;
OC Alphaherpesvirinae; Simplexvirus.
OX NCBI_TaxID=10315;
RN [1]
RP SEQUENCE FROM N.A.
RL Dolan A.;
RU Submitted (FEB-1997) to the EMBL/GenBank/DBJ databases.
CC -1- FUNCTION: PROBABLY INVOLVED IN DNA REPLICATION.
CC -1- SIMILARITY: BELONGS TO FAMILY THAT GROUPS TOGETHER HSV-1 UL8,
CC EHV-1 54, VZV 52 AND HCMV 102.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL: Z86099; CAB06768.1; -
DR InterPro: IPR004996; Herpes_HEPA.
DR Pfam: PF03324; Herpes_HEPA_1.
DR KW
SQ SEQUENCE 752 AA; 80026 MW; 8749B92360B58AD5 CRC64;
Query Match 87.6%; Score 92; DB 1; Length 752;
Best Local Similarity 85.0%; Pred. No. 3.3e-06;
Matches 17; Conservative 1; Mismatches 2; Indels 0; Gaps 0;
OY 1 IEIVFTGYLAGVWGEGSFV 20
DB 721 IQLVFTGYLAGVWGEGSFV 740
RESULT 3
YCDZ_SALT STANDARD; PRT; 160 AA.
ID YCDZ_SALT
AC O54390;
DT 15-DEC-1998 (Rel. 37, Created)
DT 28-FEB-2003 (Rel. 41, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Hypothetical protein ycdz.
GN YCDZ OR STM139.
OS Salmoneella typhimurium.
OC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
OC Enterobacteriaceae; Salmoneella.
OX NCBI_TaxID=602;
RN [1]
RP SEQUENCE FROM N.A.
RL STRAIN=LT2 / SSGC1412 / ATCC 700720;
RC MEDLINE=21534948; PubMed=11677609;
RA McCelland M., Sanderson K.E., Spieth J., Clifton S.W., Latreille P.,
RA Courtney L., Portolillo S., Ali J., Dante M., Du F., Hou S., Layman D.,
RA Leonard S., Nguyen C., Scott K., Holmes A., Grewal N., Mulvaney E.,
RA Ryan E., Sun H., Flores L., Miller W., Stoneking T., Nan M.,
RA Waterston R., Wilson R.K.;
RT "Complete genome sequence of Salmoneella enterica serovar Typhimurium
RT LT2.";
RL Nature 413:852-856(2001).
RP SEQUENCE OF 55-160 FROM N.A.

RC STRAIN=SR-11;
RX MEDLINE=98117058; PubMed=9457880;
RA Romling U., Bian Z., Hammar M., Sierraalta M.D., Normark S.;
RT "Curli fibers are highly conserved between Salmoneella typhimurium and
RT Escherichia coli with respect to operon structure and regulation.";
RL J. Bacteriol. 180:722-731(1998).
CC -1- SUBCELLULAR LOCATION: Integral membrane protein (potentia1).
CC -1- SIMILARITY: TO E COLI VAC.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL: AE008749; AL20068.1; ALT_INIT.
DR EMBL: AU002301; CA05311.1; -
DR StyGene; SGI0739; ycdz.
KW Hypothetical protein; Transmembrane; Complete proteome.
FT TRANSMEM 20 42 POTENTIAL.
FT TRANSMEM 50 70 POTENTIAL.
FT TRANSMEM 72 92 POTENTIAL.
FT TRANSMEM 99 119 POTENTIAL.
FT TRANSMEM 123 143 POTENTIAL.
SQ SEQUENCE 160 AA; 16819 MW; 5E103D7EDC6BA7DA CRC64;
Query Match 46.7%; Score 49; DB 1; Length 160;
Best Local Similarity 53.8%; Pred. No. 2.1;
Matches 7; Conservative 4; Mismatches 2; Indels 0; Gaps 0;
OY 4 VFTGYLAGVWGEG 16
DB 9 ITTGILSGIWG 21
RESULT 4
YCDZ_ECOLI STANDARD; PRT; 163 AA.
ID YCDZ_ECOLI
AC P75916;
DT 01-NOV-1997 (Rel. 35, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Hypothetical protein ycdz.
GN YCDZ OR B1036.
OS Escherichia coli.
OC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
OC Enterobacteriaceae; Escherichia.
OX NCBI_TaxID=562;
RN [1]
RP SEQUENCE FROM N.A.
RL STRAIN=K12 / MG1655;
RC MEDLINE=97426617; PubMed=9278503;
RA Blactner F.R., Plunkett G. III, Bloch C.A., Perna N.T., Burland V.,
RA Riley M., Collado-VIDES J., Glasner J.D., Rode C.K., Mayhew G.P.,
RA Gregor J., Davis N.W., Klapatrick H.A., Goeden M.A., Rose D.J.,
RA Mau B., Shao Y.;
RT "The complete genome sequence of Escherichia coli K-12.";
RL Science 277:1453-1474(1997).
RN [2]
RP SEQUENCE FROM N.A.
RL STRAIN=K12;
RC MEDLINE=97061202; PubMed=8905232;
RA Ikemoto K., Inada T., Baba T., Fujita K., Hayashi K., Honjo A.,
RA Ikemoto K., Inada T., Itoh T., Kajinara M., Kanai K., Kashimoto K.,
RA Kimura S., Kitagawa M., Makino K., Masuda S., Miki T., Mizobuchi K.,
RA Mori H., Motomura K., Nakamura Y., Nishimoto H., Nishio Y., Saito N.,
RA Sempel G., Seki Y., Tagami H., Takemoto K., Wada C., Yamamoto Y.,
RA Yano M., Horiuchi T.;
RT "A 718-kb DNA sequence of the Escherichia coli K-12 genome
RT corresponding to the 12.7-28.0 min region on the linkage map.";
RL DNA Res. 3:137-155(1996).

```

CC -1- SUBCELLULAR LOCATION: Integral membrane protein (Potential).
CC -1- SIMILARITY: TO E.COLI YAC.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; AEO00205; AAC74120.1; ALT INIT.
DR EMBL; D90740; BAA35817.1; ALT INIT.
DR EMBL; D90741; BAA35826.1; ALT_INIT.
DR Ecogene; EGI3872; ycdZ.
DR Hypothetical protein; transmembrane; Complete proteome.
KW TRANSMEM 3 23 POTENTIAL.
FT TRANSMEM 25 45 POTENTIAL.
FT TRANSMEM 49 69 POTENTIAL.
FT TRANSMEM 72 92 POTENTIAL.
FT TRANSMEM 99 119 POTENTIAL.
FT TRANSMEM 123 143 POTENTIAL.
SQ SEQUENCE 163 AA; 17074 MW; 999BB93ALFB35B68 CRC64;

Query Match 43.8%; Score 46; DB 1; Length 163;
Best Local Similarity 54.5%; Pred. No. 6.2;
Matches 6; Conservative 4; Mismatches 1; Indels 0; Gaps 0;

QY 4 VFTGVLAWG 14
Db 9 ITTGILSGIWG 19

RESULT 5
CLCB_SALTI STANDARD; PRT; 417 AA.
AC Q826Y0;
DT 28-FEB-2003 (Rel. 41, Created)
DT 10-OCT-2003 (Rel. 42, Last sequence update)
DT 10-OCT-2003 (Rel. 42, Last annotation update)
DE Voltage-gated ClC-type chloride channel CLCB.
GN CLCB OR STY1574 OR T1411.
OS Salmoneilla typhi.
OC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
OC Enterobacteriaceae; Salmoneilla.
CX NCBI_TaxID=601;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=CT18;
RX MEDLINE=21534947; PubMed=11677608;
RA Parhill J., Dougan G., James K.D., Thomson N.R., Pickard D., Main J.,
RA Churcher C., Mungall K.L., Bentley S.D., Holden M.T.G., Sebahia M.,
RA Baker S., Basham D., Brooks K., Chillingworth T., Connor P.,
RA Cronin A., Davis P., Davies R.M., Dowd I., White N., Farrar J.,
RA Feltwell T., Hamlin N., Haque A., Hien T.T., Holroyd S., Jasegla K.,
RA Krogh A., Larsen T.S., Leachter S., Moule S., O'Gaora P., Parry C.,
RA Quail M.A., Rutherford K., Simmonds W., Skelton J., Stevens K.,
RA Whitehead S., Barrett B.G.;
RT "Complete genome sequence of a multiple drug resistant Salmoneilla
RT enterica serovar Typhi CT18."
RU Nature 413:848-852(2001).
[2]
RP SEQUENCE FROM N.A.
RC STRAIN=Ty2 / ATCC 700931;
RX MEDLINE=22531367; PubMed=12644504;
RA Deng W., Itoh S.-R., Plunkett G. III, Mayhew G.F., Rose D.J.,
RA Burland V., Kodoyanni V., Schwartz D.C., Blattner F.R.;
RT "Comparative genomics of Salmoneilla enterica serovar Typhi strains Ty2
RT and CT18."
RL J. Bacteriol. 185:2330-2337(2003).
CC -1- FUNCTION: Probably acts as an electrical shunt for an outwardly-
CC directed proton pump that is linked to amino acid decarboxylation,
CC as part of the extreme acid resistance (XAR) response.

```

```

CC -1- SUBCELLULAR LOCATION: Integral membrane protein. Inner membrane
CC (Probable).
CC -1- MISCELLANEOUS: The two ClC channels in this bacterium, clcA and
CC clcC, act redundantly.
CC -1- SIMILARITY: Belongs to the chloride channel family.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; AL627270; CAD01823.1; ALT INIT.
DR EMBL; AEO16839; AAO69055.1; ALT_INIT.
DR HAMAP; MF_01203; -; 1.
DR InterPro; IPR001807; Cl-channel_volt.
DR Pfam; PF00654; Voltage_CLC; 1.
DR PRINTS; PR00762; CLCHANNEL.
KW Chloride channel; Ion transport; Ionic channel; Voltage-gated channel;
KW Chloride channel; Chloride; Inner membrane; Transmembrane;
KW Complete proteome.
FT TRANSMEM 5 27 POTENTIAL.
FT TRANSMEM 54 71 POTENTIAL.
FT TRANSMEM 146 168 POTENTIAL.
FT TRANSMEM 173 195 POTENTIAL.
FT TRANSMEM 222 244 POTENTIAL.
FT TRANSMEM 259 281 POTENTIAL.
FT TRANSMEM 288 310 POTENTIAL.
FT TRANSMEM 320 342 POTENTIAL.
FT TRANSMEM 349 371 POTENTIAL.
FT TRANSMEM 386 404 POTENTIAL.
SQ SEQUENCE 417 AA; 44200 MW; B43C8AD1268406AD CRC64;

Query Match 43.8%; Score 46; DB 1; Length 417;
Best Local Similarity 50.0%; Pred. No. 15;
Matches 7; Conservative 5; Mismatches 2; Indels 0; Gaps 0;

QY 1 IEVFTGVLAWG 14
Db 222 VMIVSTGLVGLWG 235

RESULT 6
OL56_STRAT STANDARD; PRT; 3519 AA.
ID OL56_STRAT
AC Q07017;
DT 01-NOV-1997 (Rel. 35, Created)
DT 01-NOV-1997 (Rel. 35, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Oleandomycin polyketide synthase, modules 5 and 6.
GN ORB.
OS Streptomyces antibioticus.
OC Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;
OC Streptomycineae; Streptomycetaceae; Streptomyces.
CX NCBI_TaxID=1890;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=94150470; PubMed=8107683;
RA Swan D.G., Rodriguez A.M., Vilches C., Mendez C., Salas J.A.;
RT "Characterisation of a Streptomyces antibioticus gene encoding a type
RT I polyketide synthase which has an unusual coding sequence."
RL Mol. Genet. 242:358-362(1994).
CC -1- FUNCTION: MAY BE INVOLVED IN THE BIOSYNTHESIS OF THE OLEANDOMYCIN
CC LACTONE RING.
CC -1- COFACTOR: Contains 2 covalently bound phosphopantetheines.
CC -1- SIMILARITY: Contains 2 acyl carrier domains.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way

```

CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (see <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC
 DR EMBL; L09654; AAA19695.1; -
 DR PIR; S43048; S43048.
 DR HSP; P25715; 1MLA.
 DR InterPro; IPR001227; Ac trans.
 DR InterPro; IPR000794; ketoacyl_synth.
 DR InterPro; IPR006163; Pp bind.
 DR InterPro; IPR006162; Pp bind.
 DR InterPro; IPR001031; Thioesterase.
 DR Pfam; PF00698; Acyl_transf. 2.
 DR Pfam; PF00109; ketoacyl-synt. 2.
 DR Pfam; PF02801; ketoacyl-synt C. 2.
 DR Pfam; PF00550; pp-binding; 2.
 DR Pfam; PF00975; Thioesterase; 1.
 DR PROSITE; PS00012; PHOSPHOPANTHETINE; 2.
 DR PROSITE; PS00806; B_KETOACYL_SYNTHASE; 2.
 DR PROSITE; PS50075; ACP DOMAIN; 2.
 DR Transferrase; Acyltransferase; Antibiotic biosynthesis; NADP;
 KW Phosphopantetheine; Multifunctional enzyme; Repeat.
 FT DOMAIN 1
 FT DOMAIN 2 3519
 FT DOMAIN 3 501
 FT DOMAIN 4 569
 FT DOMAIN 5 1200
 FT DOMAIN 6 1487
 FT DOMAIN 7 1561
 FT DOMAIN 8 1686
 FT DOMAIN 9 2220
 FT DOMAIN 10 2856
 FT DOMAIN 11 3141
 FT DOMAIN 12 3270
 FT DOMAIN 13 3519
 FT ACT SITE 210
 FT NP BIND 1203
 FT BINDING 1524
 FT ACT SITE 1859
 FT NP BIND 2311
 FT BINDING 2859
 FT BINDING 3178
 FT SEQUENCE 3519 AA; 368561 MW; 41AE78AAAE61F86 CRC64;
 Query Match 43.8%; Score 46; DB 1; Length 3519;
 Best Local Similarity 72.7%; Pred. No. 1.3e+02;
 Matches 8; Conservative 1; Mismatches 2; Indels 0; Gaps 0;
 QY 10 AGWGGGKGFV 20
 DB 1337 AGWGGGQAV 1347
 RESULT 7
 ENV_HV1KB STANDARD; PRT; 861 AA.
 ID ENV_HV1KB
 AC P31819;
 DT 01-JUL-1993 (Rel. 26, Created)
 DT 01-JUL-1993 (Rel. 26, Last sequence update)
 DT 15-JUL-1999 (Rel. 38, Last annotation update)
 DE Erythrocyte polypeptide GP160 precursor (contains: Exterior membrane
 glycoprotein (GP120); Transmembrane glycoprotein (GP41)).
 GN ENV.
 OS Human immunodeficiency virus type 1 (KB-1 isolate) (HIV-1).
 OC Viruses; Retroviral viruses; Retroviridae; Lentivirus.
 OX NCBI_TaxID=36375;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=92351552; Pubmed=1322587;
 RA Shimizu H., Hasebe F., Tsuchie H., Morikawa S., Ushijima H.,
 RA Kitamura T.;
 RT "Analysis of a human immunodeficiency virus type 1 isolate carrying a
 RT truncated transmembrane glycoprotein.",
 RL Virology 189:534-546 (1992).

CC -1- MISCELLANEOUS: READTHROUGH OF THE TERMINATOR UAG OCCURS BETWEEN
 CC THE COONS FOR 729-ALA AND 730-ARG.
 CC
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (see <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC
 DR EMBL; D12582; BAA02124.1; ALT_SEQ.
 DR PIR; A42995; VCLUBK.
 DR InterPro; IPR000328; Env GP41.
 DR InterPro; IPR000777; GP120.
 DR Pfam; PF00516; GP120. 1.
 DR Pfam; PF00517; GP41. 1.
 DR Aids; Coat protein; Polyprotein; Glycoprotein; Transmembrane;
 KW Aids; Signal.
 FT SIGNAL 1 35
 FT CHAIN 36 517
 FT CHAIN 518 861
 FT TRANSMEM 690 711
 FT DISULFID 59 79
 FT DISULFID 124 212
 FT DISULFID 131 203
 FT DISULFID 136 160
 FT DISULFID 225 254
 FT DISULFID 235 246
 FT DISULFID 303 337
 FT DISULFID 383 446
 FT DISULFID 380 419
 FT CARBOHYD 93 93
 FT CARBOHYD 141 141
 FT CARBOHYD 145 145
 FT CARBOHYD 146 146
 FT CARBOHYD 159 159
 FT CARBOHYD 163 163
 FT CARBOHYD 191 191
 FT CARBOHYD 192 192
 FT CARBOHYD 237 237
 FT CARBOHYD 241 241
 FT CARBOHYD 248 248
 FT CARBOHYD 269 269
 FT CARBOHYD 283 283
 FT CARBOHYD 296 296
 FT CARBOHYD 308 308
 FT CARBOHYD 338 338
 FT CARBOHYD 345 345
 FT CARBOHYD 361 361
 FT CARBOHYD 367 367
 FT CARBOHYD 397 397
 FT CARBOHYD 403 403
 FT CARBOHYD 408 408
 FT CARBOHYD 414 414
 FT CARBOHYD 449 449
 FT CARBOHYD 465 465
 FT CARBOHYD 468 468
 FT CARBOHYD 617 617
 FT CARBOHYD 622 622
 FT CARBOHYD 631 631
 FT CARBOHYD 643 643
 FT CARBOHYD 821 821
 FT SEQUENCE 861 AA; 98116 MW; 30C6787658F0C9DA CRC64;
 Query Match 42.9%; Score 45; DB 1; Length 861;
 Best Local Similarity 58.3%; Pred. No. 45;
 Matches 7; Conservative 2; Mismatches 3; Indels 0; Gaps 0;
 QY 9 LAGWGGGKGFV 20
 DB 598 LMGWGGGKRTI 609

```

RESULT 8
DIT1_YEAST
ID DIT1_YEAST STANDARD; PRT; 536 AA.
AC P21623;
DT 01-MAY-1991 (Rel. 18, Created)
DT 01-OCT-1996 (Rel. 34, Last sequence update)
DT 15-DEC-1998 (Rel. 37, Last annotation update)
DE Spore wall maturation protein DIT1.
DI DIT1 OR YDR403W OR D9509.21.
OS Saccharomyces cerevisiae (Baker's Yeast).
OC Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
  Saccharomycetales; Saccharomycetaceae; Saccharomycetes.
OX NCBI_TaxID=4932;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=AB320;
RX MEDLINE=91065523; PubMed=2249774;
RA Brita P., Breitenbach M., Ellinger A., Segall J.;
  "Isolation of two developmentally regulated genes involved in spore
  wall maturation in Saccharomyces cerevisiae.";
  Genes Dev. 4:1775-1789(1990).
RL [2]
RN [3]
RP SEQUENCE FROM N.A.
RA Dietrich F.S., Mulligan J., Allen E., Araujo R., Aviles E.,
  Berro A., Carpenter J., Chen E., Cherry J.M., Chung E., Duncan M.,
  Hunkle-Smith S., Hyman R., Komp C., Lashkari D., Lew H., Lin D.,
  Mosedale D., Nakahara K., Namath A., Oetner P., Oh C., Petel F.X.,
  Roberts D., Schramm S., Schroeder M., Shogen T., Shroff N.,
  Winant A., Yelton M., Botstein D., Davis R.W.;
  Submitted (Aug-1995) to the EMBL/GenBank/DBJ databases.
  [3]
RP CHARACTERIZATION.
RX MEDLINE=94240168; PubMed=8183942;
RA Brita P., Eckersdorfer M., Breitenbach M.;
  "The sporulation-specific enzymes encoded by the DIT1 and DIT2 genes
  catalyze a two-step reaction leading to a soluble Lt-dityrosine-
  containing precursor of the yeast spore wall.";
  Proc. Natl. Acad. Sci. U.S.A. 91:4524-4528(1994).
RL [1]
RP FUNCTION: Involved in spore wall maturation. Catalyzes a two step
  reaction that leads to the Lt-dityrosine containing precursor of
  the spore wall.
CC -!- DEVELOPMENTAL STAGE: Sporulation.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
  between the Swiss Institute of Bioinformatics and the EMBL outstation -
  the European Bioinformatics Institute. There are no restrictions on its
  use by non-profit institutions as long as its content is in no way
  modified and this statement is not removed. Usage by and for commercial
  entities requires a license agreement (See http://www.isb-sib.ch/announce/
  or send an email to license@isb-sib.ch).
  -----
CC EMBL: X55712; CAA39245.1; -
CC EMBL: U32274; AAB6485.1; -
CC PIR: A36395; A36395.
CC Germonline: 140895; -
CC SGD: S0002811; DIT1.
CC GO: GO:0003824; F:catalytic activity; IDA.
CC GO: GO:0007152; P:spore wall assembly (sensu Saccharomycetes); IDA.
CC InterPro: IPR007817; DIT1_PVCA.
CC Pfam: PF05141; DIT1_PVCA_1.
CC Spoolation. 273
CC CONFLICT 273 F -> S (IN REF. 1).
CC FT
CC SEQUENCE 536 AA; 61390 MW; 4E7C8AC0FD4EEDD8 CRC64;
  SO
  Query Match 42.4%; Score 44.5; DB 1; Length 536;
  Best Local Similarity 42.3%; Pred. No. 33;
  Matches 11; Conservative 2; Mismatches 6; Indels 7; Gaps 1;
  QY 2 ELVFTGVLAGVW-----GGGGRKFV 20
  DB 506 FALKKMGVGVWVWVTRFDIGGGHFV 531

```

```

RESULT 9
FABG_MYCSM
ID FABG_MYCSM STANDARD; PRT; 255 AA.
AC P71534;
DT 01-NOV-1997 (Rel. 35, Created)
DT 01-NOV-1997 (Rel. 35, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE 3-oxoacyl-[acyl-carrier protein] reductase (SC 1.1.1.100) (3-ketoacyl-
  acyl carrier protein reductase).
DI M0441.
OS Mycobacterium smegmatis.
OC Bacteria; Actinobacteria; Actinomycetidae; Actinomycetales;
  Corynebacterineae; Mycobacteriaceae; Mycobacterium.
OX NCBI_TaxID=1772;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=ATCC 700084 / mc(2)155;
RA Banerjee A., Sugantino M., Sacchettini J.C., Jacobs W.R. Jr.;
  "Molecular cloning, expression and characterization of 3-ketoacyl
  reductase from mycobacteria.";
  Submitted (OCT-1996) to the EMBL/GenBank/DBJ databases.
RL [1]
RN [2]
RP CATALYTIC ACTIVITY: (3R)-3-hydroxyacyl-[acyl-carrier protein] +
  NADP(+) = 3-oxoacyl-[acyl-carrier protein] + NADPH.
CC -!- PATHWAY: Fatty acid biosynthesis pathway; first reduction step.
CC -!- SIMILARITY: Belongs to the short-chain dehydrogenases/reductases
  (SDR) family.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
  between the Swiss Institute of Bioinformatics and the EMBL outstation -
  the European Bioinformatics Institute. There are no restrictions on its
  use by non-profit institutions as long as its content is in no way
  modified and this statement is not removed. Usage by and for commercial
  entities requires a license agreement (See http://www.isb-sib.ch/announce/
  or send an email to license@isb-sib.ch).
  -----
CC EMBL: U66800; AAC69638.1; -
CC HSP: P47227; 1BD.
CC InterPro: IPR002198; ADH short.
CC Pfam: PF00106; adh_short_1.
CC PRINTS: PR00080; SDRFAMILY.
CC PROSITE: PS00061; ADH_SHORT, 1.
CC K1 Fatty acid biosynthesis; Oxidoreductase; NADP.
CC FT AC SITE 10 39 NADP (BY SIMILARITY).
CC AC SITE 161 161 BY SIMILARITY.
CC SEQUENCE 255 AA; 26769 MW; DD76495B0A8BBD6 CRC64;
  SO
  Query Match 41.9%; Score 44; DB 1; Length 255;
  Best Local Similarity 37.5%; Pred. No. 19;
  Matches 6; Conservative 6; Mismatches 4; Indels 0; Gaps 0;
  QY 3 LVFTGVLAGVWGGGK 18
  DB 143 IIFGVSVMGIGIQ 158
  RESULT 10
  Y441_METUA
  ID Y441_METUA STANDARD; PRT; 267 AA.
  AC O57883;
  DT 01-NOV-1997 (Rel. 35, Created)
  DT 01-NOV-1997 (Rel. 35, Last sequence update)
  DT 16-OCT-2001 (Rel. 40, Last annotation update)
  DE Hypothetical protein M0441.
  DI M0441.
  OS Methanococcus jannaschii.
  OC Archaea; Euryarchaeota; Methanococci; Methanococcales;
  OC Methanocaldococcaceae; Methanocaldococcus.
  OX NCBI_TaxID=2190;
  RN [1]
  RP SEQUENCE FROM N.A.
  RC STRAIN=JAL-1 / DSM 2661 / ATCC 43067;
  RX MEDLINE=96337999; PubMed=8688087;

```

RA Bult C.J., White O., Olsen G.J., Zhou L., Fleischmann R.D.,
 RA Sutton G.G., Blake J.A., Fitzgerald L.M., Clayton R.A., Gocayne J.D.,
 RA Kerlavage A.R., Dougherty B.A., Tomb J.F., Adams M.D., Reich C.I.,
 RA Kierstead R., Kirkness E.F., Weinstock K.G., Merriam J.M., Glodek A.,
 RA Scott J.L., Geoghegan N.S.M., Weidman J.F., Fuhrmann J.L., Nguyen D.,
 RA Otterbach T.R., Kelley J.M., Peterson J.D., Sadow P.W., Hanna M.C.,
 RA Cotton M.D., Roberts K.M., Hurt M.A., Kaine B.P., Borodovsky M.,
 RA Klenk H.-P., Fraser C.M., Smith H.O., Woese C.R., Venter J.C.;
 RA "Complete genome sequence of the methanogenic archaeon, *Methanococcus*
 RT *jannaschii*.";
 RT Science 273:1058-1073 (1996).
 CC -1- SUBCELLULAR LOCATION: Integral membrane protein (Potential).
 CC
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC
 CC EMBL: U67495; AAB98428.1; -;
 DR PIR: A64355; A64355.
 DR TIGR: MJ0441; -;
 DR InterPro: IPR002781; DUF81.
 DR Pfam: PF01925; DUF81; 1.
 KM Hypothetical protein; Transmembrane; Complete proteome.
 FT TRANSMEM 10 30 POTENTIAL.
 FT TRANSMEM 31 51 POTENTIAL.
 FT TRANSMEM 55 75 POTENTIAL.
 FT TRANSMEM 87 107 POTENTIAL.
 FT TRANSMEM 158 178 POTENTIAL.
 FT TRANSMEM 185 205 POTENTIAL.
 FT TRANSMEM 213 233 POTENTIAL.
 SQ SEQUENCE 267 AA; 28980 MM; 08EF7F36D198FC60 CRC64;
 Query Match 41.9%; Score 44; DB 1; Length 267;
 Best Local Similarity 52.9%; Pred. No. 20;
 Matches 9; Conservative 3; Mismatches 5; Indels 0; Gaps 0;
 QY 4 VFTGVLAGVWGEGKRFV 20
 Db 154 VTGFLSGFSGIGGIV 170
 RESULT 11
 GCSA_CHLITE STANDARD; PRT: 444 AA.
 ID Q8K05; -;
 AC 10-OCT-2003 (Rel. 42, Created)
 DT 10-OCT-2003 (Rel. 42, Last sequence update)
 DT 10-OCT-2003 (Rel. 42, Last annotation update)
 DE Probable glycine dehydrogenase [decarboxylating] subunit 1
 DE (EC 1.4.4.2) (Glycine decarboxylase subunit 1) (Glycine cleavage
 DE system P-protein subunit 1).
 GN GCVPA OR GCVPI OR CT1625.
 OS Chlorobium tepidum.
 OC Bacteria; Chlorobi; Chlorobia; Chlorobiales; Chlorobiaceae;
 OC Chlorobiium.
 OC NCBI_TaxID=1097;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=TL5 / ATCC 49652 / DSM 12025;
 RX MEDLINE=22103685; PubMed=12093901;
 RA Eisen J.A., Nelson K.E., Paulsen I.T., Heidelberg J.F., Wu M.,
 RA Dodson R.J., Deboy R., Gwinn M.L., Nelson W.C., Haft D.H.,
 RA Hickey E.K., Peterson J.D., Durkin A.S., Kolonay J.L., Yang F.,
 RA Holt I., Umayam L.A., Mason T., Brenner M., Shea T.P., Parksey D.,
 RA Niernan W.C., Feldblyum T.V., Hansen C.L., Craven M.B., Radue D.,
 RA Vanachavan J., Kouri H., White O., Gruber T.W., Ketchum K.A.,
 RA Venter J.C., Tettelin H., Bryant D.A., Fraser C.M.;
 RA "The complete genome sequence of *Chlorobium tepidum* TL5, a
 RT photosynthetic, anaerobic, green-sulfur bacterium.";

RL Proc. Natl. Acad. Sci. U.S.A. 99:9509-9514(2002).
 CC -1- FUNCTION: The glycine cleavage system catalyzes the degradation of
 CC glycine. The P protein binds the alpha-amino group of glycine
 CC through its pyridoxal phosphate cofactor; CO(2) is released and
 CC the remaining methylene moiety is then transferred to the
 CC lipamide cofactor of the H protein (By similarity).
 CC -1- CATALYTIC ACTIVITY: Glycine + lipoylprotein = S-
 CC aminomethylidene-diolipoylprotein + CO(2).
 CC -1- SUBUNIT: The glycine cleavage system is composed of four proteins:
 CC P, T, L and H. In this organism, the P 'protein' is an heterodimer
 CC of two subunits (By similarity).
 CC -1- SIMILARITY: Belongs to the gcvP family. N-terminal subunit
 CC subfamily.
 CC
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC
 CC EMBL: AE012917; AAM72850.1; -;
 DR TIGR: CT1625; -;
 DR HAMAP: MF_00712; -; 1.
 DR InterPro: IPR003437; GDC-P.
 DR Pfam: PF02347; GDC-P; 1.
 KM Oxidoreductase; Complete proteome.
 SQ SEQUENCE 444 AA; 47739 MM; 334273FF9723B87F CRC64;
 Query Match 41.9%; Score 44; DB 1; Length 444;
 Best Local Similarity 52.9%; Pred. No. 33;
 Matches 9; Conservative 0; Mismatches 8; Indels 0; Gaps 0;
 QY 4 VFTGVLAGVWGEGKRFV 20
 Db 406 VFAVDLSAMSGDGLV 422
 RESULT 12
 PMPD_CHLMU STANDARD; PRT: 1520 AA.
 ID Q9PLE0; -;
 AC 16-OCT-2001 (Rel. 40, Created)
 DT 16-OCT-2001 (Rel. 40, Last sequence update)
 DT 16-OCT-2001 (Rel. 40, Last annotation update)
 DE Probable outer membrane protein pmpD precursor (Polymorphic membrane
 DE protein D).
 GN PMPD OR TC0197.
 OS Chlamydia muridarum.
 OC Bacteria; Chlamydiae; Chlamydiales; Chlamydiaceae; Chlamydia.
 OC NCBI_TaxID=83560;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=Mopn / Ni99;
 RX MEDLINE=20150255; PubMed=10684935.
 RA Read T.D., Brunnham R.C., Shen C., Gill S.R., Heidelberg J.F.,
 RA White O., Hickey E.K., Peterson J., Utterback T., Berry K., Baas S.,
 RA Linner K., Weidman J., Kouri H., Craven M., Bowman C., Dodson R.,
 RA Gwinn M., Nelson W., Deboy R., Kolonay J., McClarty G., Salzberg S.L.,
 RA Eisen J., Fraser C.M.;
 RA "Genome sequences of *Chlamydia trachomatis* Mopn and *Chlamydia*
 RT *pneumoniae* AR39.";
 RT Nucleic Acids Res. 28:1397-1406 (2000).
 RL -1- SUBCELLULAR LOCATION: Cell wall surface (elementary bodies)
 CC (Potential).
 CC -1- SIMILARITY: Belongs to the PMP outer membrane protein family.
 CC
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial


```

FT CARBOHYD 319 319 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 374 374 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 380 380 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 463 463 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 474 474 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 479 479 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 628 628 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 637 637 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 653 653 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 762 762 N-LINKED (GLCNAC. . .) (POTENTIAL).
SQ SEQUENCE 682 AA; 101196 MW; 216DD5C06C805DA7 CRC64;

```

```

Query Match 41.4%; Score 43.5; DB 1; Length 682;
Best Local Similarity 40.9%; Pred. No. 77;
Matches 9; Conservative 3; Mismatches 3; Indels 7; Gaps 1;

```

```

OY 6 TGVLAGVWGE-----GKRV 20
DB 845 TETLAGAGHLEWALQRCGRWI 866

```

RESULT 15

```

YA00_PYRHO STANDARD; PRT; 355 AA.
AC OS8728;
DT 16-OCT-2001 (Rel. 40, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Hypothetical protein PH1000.
GN PH1000.
OS Pyrococcus horikoshii.
OC Archaea; Euryarchaeota; Thermococci; Thermococcales; Thermococcaceae;
OC Pyrococcus.
OX NCBI_TaxID=53953;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=OT3;
RX MEDLINE=98344137; PubMed=9679194;
RA Kawarabayashi Y., Sawada M., Horikawa H., Haikawa Y., Hino Y.,
RA Yamamoto S., Sekine M., Baba S.-I., Kosugi H., Hosoyama A., Nagai Y.,
RA Sakai M., Ogura K., Otsuka R., Nakazawa M., Takamaya M., Ohfuku Y.,
RA Funahashi T., Tanaka T., Kudo Y., Yamazaki J., Kishida N., Oguchi A.,
RA Aoki K.-I., Yoshizawa T., Nakamura Y., Robb F.T., Horikoshi K.,
RA Masuchi Y., Shizuya H., Kikuchi H.;
RT "Complete sequence and gene organization of the genome of a hyper-
thermophilic archaeobacterium, Pyrococcus horikoshii OT3.";
RL DNA Res. 5:55-76(1998).
CC -!- SUBCELLULAR LOCATION: Integral membrane protein (Potential).
CC -!- SIMILARITY: BELONGS TO THE UPF0118 (PERM) FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL: AP000004; BAA30097.1; -
DR PIR: C71092; C71092.
DR InterPro: IPR002549; UPF0118.
DR Pfam: PF01594; UPF0118; 1.
KW Hypothetical protein; Transmembrane; Complete proteome.
FT TRANSMEM 34 54 POTENTIAL.
FT TRANSMEM 55 75 POTENTIAL.
FT TRANSMEM 84 104 POTENTIAL.
FT TRANSMEM 158 178 POTENTIAL.
FT TRANSMEM 212 232 POTENTIAL.
FT TRANSMEM 240 260 POTENTIAL.
FT TRANSMEM 274 294 POTENTIAL.
FT TRANSMEM 310 330 POTENTIAL.
SQ SEQUENCE 355 AA; 40179 MW; D0929BBFEA847F39 CRC64;

```

```

Query Match 41.0%; Score 43; DB 1; Length 355;
Best Local Similarity 50.0%; Pred. No. 37;
Matches 8; Conservative 2; Mismatches 6; Indels 0; Gaps 0;

```

```

OY 3 LVFTGLAGVWGEQK 18
DB 311 LVLTGMIGTWTGLGLX 326

```

```

Search completed: September 30, 2004, 11:10:31
Job time : 8.42857 secs

```

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: September 30, 2004, 10:58:43 ; Search time 12.7619 Seconds
(without alignments)
150.748 Million cell updates/sec

Title: US-09-988-851A-7
Perfect score: 105
Sequence: 1 IEIVFTGVLAVGWGEGKRFV 20

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 283366 seqs, 96191526 residues

Total number of hits satisfying chosen parameters: 283366

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%

Listing first 45 summaries

Database :
1: PIR.78:*
2: PIR1:*
3: PIR3:*
4: PIR4:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	105	100.0	750	1 WMBEX8	U18 protein - huma
2	55	52.4	226	2 C70383	conserved hypotet
3	55	49.5	255	2 B64604	conserved hypotet
4	52	49.5	255	2 D71910	hypothetical prote
5	52	49.5	721	2 C84677	probable membrane
6	50	47.6	328	2 C87673	4-hydroxybenzoate
7	49	46.7	251	2 T44070	conserved hypotet
8	47	44.8	261	2 H69217	hypothetical prote
9	47	44.8	651	2 C69374	conserved hypotet
10	46	43.8	170	2 G82446	conserved hypotet
11	46	43.8	176	2 AC0635	probable membrane
12	46	43.8	179	2 A64846	probable membrane
13	46	43.8	179	2 B90805	hypothetical prote
14	46	43.8	179	2 B85665	hypothetical prote
15	46	43.8	429	2 AF0681	probable voltage s
16	46	43.8	703	2 T05632	hypothetical prote
17	46	43.8	3519	2 S43048	polyketide synthas
18	45.5	43.3	512	2 S75887	hypothetical prote
19	45.5	43.3	4735	2 T17463	rifamycin polypepti
20	45	42.9	191	2 B72480	hypothetical prote
21	45	42.9	277	2 B64548	conserved hypotet
22	45	42.9	277	2 F71960	hypothetical prote
23	45	42.9	293	2 T50940	D1d protein limpo
24	45	42.9	475	2 G69265	hypothetical prote
25	45	42.9	861	2 VCLJXX	env polypeptide pr
26	45	42.9	729	2 VCLJXX	env polypeptide pr
27	44.5	42.4	536	2 A36395	spore wall maturat
28	44	41.9	115	2 F72569	hypothetical prote
29	44	41.9	214	2 T44107	conserved hypotet

30	44	41.9	214	2 G89763	conserved hypotet
31	44	41.9	252	2 A71165	hypothetical prote
32	44	41.9	253	2 F75067	hypothetical prote
33	44	41.9	267	2 A64355	hypothetical prote
34	44	41.9	293	2 D72410	prolipoprotein dia
35	44	41.9	308	2 H90452	conserved hypotet
36	44	41.9	401	2 AD1264	ammonium transport
37	44	41.9	401	2 AF1626	ammonium transport
38	44	41.9	480	2 C69438	hypothetical prote
39	44	41.9	533	2 S47582	high-affinity pota
40	44	41.9	1489	2 D70807	hypothetical glyci
41	44	41.9	1520	2 A81731	polymorphic membra
42	44	41.9	1901	2 F70906	hypothetical glyci
43	43.5	41.4	402	2 H81296	probable integral
44	43.5	41.4	881	2 VCLJG3	env polypeptide -
45	43	41.0	159	2 G84125	hypothetical prote

ALIGNMENTS

```

RESULT 1
WMBEX8
U18 protein - human herpesvirus 1 (strain 17)
C:Species: human herpesvirus 1
C>Date: 31-Dec-1989 #sequence_revision 31-Dec-1989 #text_change 16-Jun-2000
C:Accession: C29690; H28133
R:McGeoch, D.J.; Dalrymple, M.A.; Dolan, A.; McNab, D.; Perry, L.J.; Taylor, P.; Chalber
J.; Virol. 62, 444-453, 1988
A:Title: Structures of herpes simplex virus type 1 genes required for replication of viri
A:Reference number: A93040; PMID:88091053; PMID:2826807
A:Accession: C29690
A:Molecule type: DNA
A:Residues: 1-750 <MCG>
A:Cross-references: GB:M19120; NID:G330226; PIDN:AAA45823.1; PID:G330235
R:McGeoch, D.J.; Dalrymple, M.A.; Davison, A.J.; Dolan, A.; Frame, M.C.; McNab, D.; Perry
J.; Gen. Virol. 69, 1531-1574, 1988
A:Title: The complete DNA sequence of the long unique region in the genome of herpes sim
A:Reference number: A30083; PMID:88274377; PMID:2839594
A:Accession: H28133
A:Status: nucleic acid sequence not shown; translation not shown
A:Molecule type: DNA
A:Residues: 1-750 <MCG2>
A:Cross-references: GB:X14112; NID:G1944536; PIDN:CAA32344.1; PID:G59509; GB:D00317
C:Comment: This protein is required for replication of viral DNA.
C:Genetic:
A:Gene: U18
C:Superfamily: herpesvirus U18 protein
C:Keywords: DNA biosynthesis

Query Match 100.0%; Score 105; DB 1; Length 750;
Best Local Similarity 100.0%; Pred. No. 1.4e-07;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Ox 1 IEIVFTGVLAVGWGEGKRFV 20
Db 719 IEIVFTGVLAVGWGEGKRFV 738

RESULT 2
C70383
conserved hypothetical protein aq_963 - Aquifex aeolicus
C:Species: Aquifex aeolicus
C>Date: 08-May-1998 #sequence_revision 08-May-1998 #text_change 05-Nov-1999
C:Accession: C70383
R:Decker, G.; Warren, P.V.; Gaasterland, T.; Young, W.G.; Lenox, A.L.; Graham, D.E.; Ovi
V.
Nature 392, 353-358, 1998
A:Title: The complete genome of the hyperthermophilic bacterium Aquifex aeolicus.
A:Reference number: A70300; PMID:98196666; PMID:9537320
A:Accession: C70383
A:Status: preliminary; nucleic acid sequence not shown; translation not shown
A:Molecule type: DNA

```

A:Residues: 1-226 <ADF>
A:Cross-references: GB:AE000715; NID:G2983460; PIDN:AACT0057.1; PID:G2983477; GB:AE00065
A:Experimental source: strain VFS
C:Genetics:
A:Gene: aq_963

Query Match 52.4%; Score 55; DB 2; Length 226;
Best Local Similarity 50.0%; Pred. No. 1;
Matches 9; Conservative 4; Mismatches 5; Indels 0; Gaps 0;

QY 3 LVFTGLAGVWGEGGKRV 20
:||||:|
Db 163 VFTGLANWAGGNTV 180

RESULT 3
B64604
conserved hypothetical integral membrane protein HP0677 - Helicobacter pylori (strain 26
C:Species: Helicobacter pylori
C>Date: 09-Aug-1997 #sequence_revision 09-Aug-1997 #text_change 17-Mar-2000
C:Accession: B64604
R:Tom, J.F.; White, O.; Kervavage, A.R.; Clayton, R.A.; Sutton, G.G.; Fleischmann, R.D.;
Peterson, S.; Loftus, B.; Richardson, D.; Dodson, R.; Khaliq, H.G.; Glodek, A.; McKenne-
son, J.D.; Kelley, J.M.; Cotton, M.D.; Weidman, J.M.; Fujii, C.; Bowman, C.; Matthey, L.
Nature 388, 539-547, 1997
A:Authors: Wallin, E.; Hayes, W.S.; Borodovsky, M.; Karpk, P.D.; Smith, H.O.; Fraser, C.
A:Title: The complete genome sequence of the gastric pathogen Helicobacter pylori.
A:Reference number: A64520; MUID:97394467; PMID:9252185
A:Accession: B64604
A:Status: preliminary; nucleic acid sequence not shown; translation not shown
A:Molecule type: DNA
A:Residues: 1-255 <TOM>
A:Cross-references: GB:AE000581; GB:AE000511; NID:G2313802; PIDN:AA014882.1; PID:G231380
C:Superfamily: hypothetical protein HP0902

Query Match 49.5%; Score 52; DB 2; Length 255;
Best Local Similarity 47.1%; Pred. No. 3.2;
Matches 8; Conservative 6; Mismatches 3; Indels 0; Gaps 0;

QY 4 VFTGLAGVWGEGGKRV 20
:||||:|
Db 12 LFTGLSGIRFGGLI 28

RESULT 4
D71910
hypothetical protein jhp0619 - Helicobacter pylori (strain J99)
C:Species: Helicobacter pylori
A:Variety: strain J99
C>Date: 12-Feb-1999 #sequence_revision 12-Feb-1999 #text_change 17-Mar-2000
C:Accession: D71910
R:Alm, R.A.; Ling, L.S.L.; Moir, D.T.; King, B.L.; Brown, E.D.; Doig, P.C.; Smith, D.R.;
Ives, C.; Gibson, R.; Merberg, D.; Mills, S.D.; Jiang, Q.; Taylor, D.E.; Vovis, G.F.;
Nature 397, 176-180, 1999
A:Title: Genomic sequence comparison of two unrelated isolates of the human gastric path
A:Reference number: A71800; MUID:99120557; PMID:9923682
A:Accession: D71910
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-255 <ARH>
A:Cross-references: GB:AE001493; GB:AE001439; NID:G415516; PIDN:AA06192.1; PID:G415516
C:Genetics:
A:Experimental source: strain J99
A:Gene: jhp0619
C:Superfamily: hypothetical protein HP0902

Query Match 49.5%; Score 52; DB 2; Length 255;
Best Local Similarity 47.1%; Pred. No. 3.2;
Matches 8; Conservative 6; Mismatches 3; Indels 0; Gaps 0;

QY 4 VFTGLAGVWGEGGKRV 20
:||||:|
Db 12 LFTGLSGIRFGGLI 28

RESULT 5
C84677
probable membrane transporter [imported] - Arabidopsis thaliana

C:Species: Arabidopsis thaliana (mouse-ear cress)
C>Date: 02-Feb-2001 #sequence_revision 02-Feb-2001 #text_change 02-Feb-2001
C:Accession: C84677
R:Lin, X.; Kaul, S.; Rounsley, S.D.; Shea, T.P.; Benito, M.I.; Town, C.D.; Fujii, C.Y.;
M.; Xoo, H.; Moffat, K.S.; Cronin, L.A.; Shen, M.; Vankken, S.E.; Umayam, L.; Tallon, L.;
euss, D.; Nierman, W.C.; White, O.; Eisen, J.A.; Salzberg, S.L.; Fraser, C.M.; Venter, J.
Nature 402, 761-768, 1999
A:Title: Sequence and analysis of chromosome 2 of the plant Arabidopsis thaliana.
A:Reference number: A64420; MUID:20083487; PMID:10617197
A:Accession: C84677
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-721 <STO>
A:Cross-references: GB:AE002093; NID:G3860251; PIDN:AACT3019.1; GSPDB:GN00139
C:Genetics:
A:Gene: At2g7810
A:Map position: 2

Query Match 49.5%; Score 52; DB 2; Length 721;
Best Local Similarity 75.0%; Pred. No. 8.7;
Matches 9; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 5 FTGVLAVWGEG 16
:||||:|
Db 492 FTGVLAVWGEGT 503

RESULT 6
C87673
4-hydroxybenzoate octaprenyltransferase [imported] - Caulobacter crescentus
C:Species: Caulobacter crescentus
C>Date: 20-Apr-2001 #sequence_revision 20-Apr-2001 #text_change 10-May-2001
C:Accession: C87673
R:Nierman, W.C.; Feldblyum, T.V.; Paulsen, I.T.; Nelson, K.E.; Eisen, J.; Heidelberg, J.I.
B.; Laub, M.T.; DeBoy, R.T.; Dodson, R.U.; Durkin, A.S.; Gwinn, M.L.; Haft, D.H.; Kolonit-
s, J.; Ermolaeva, M.; White, O.; Salzberg, S.L.; Shapiro, L.; Venter, J.C.; Fraser, C.M.
Proc. Natl. Acad. Sci. U.S.A. 98, 4136-4141, 2001
A:Title: Complete Genome Sequence of Caulobacter crescentus.
A:Reference number: A67249; MUID:21173698; PMID:11259647
A:Accession: C87673
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-328 <STO>
A:Cross-references: GB:AE005673; NID:G13425135; PIDN:AAK25383.1; GSPDB:GN00148
C:Genetics:
A:Gene: C03421
C:Superfamily: 4-hydroxybenzoate octaprenyltransferase

Query Match 47.6%; Score 50; DB 2; Length 328;
Best Local Similarity 52.9%; Pred. No. 8;
Matches 9; Conservative 3; Mismatches 5; Indels 0; Gaps 0;

QY 3 LVFTGLAGVWGEGKRV 19
:||||:|
Db 312 LVFTGLAVGLWKFVGSF 328

RESULT 7
T44070
conserved hypothetical protein [imported] - Staphylococcus aureus
C:Species: Staphylococcus aureus
C>Date: 21-Jan-2000 #sequence_revision 21-Jan-2000 #text_change 11-May-2000
C:Accession: T44070
R:Itto, T.; Katayama, Y.; Hiratsuka, K.
Antimicrob. Agents Chemother. 43, 1449-1458, 1999
A:Title: Cloning and nucleotide sequence determination of the entire mec DNA of pre-meth
A:Reference number: Z2735; MUID:99278010; PMID:10348769
A:Accession: T44070

A>Status: preliminary; translated from GB/EMBL/DBJ
 A:Molecule type: DNA
 A:Residues: 1-251 <IT0>
 A:Cross-references: EMBL:D86934; PIDN:BA82172.1
 A:Experimental source: strain N15
 C:Superfamily: hypothetical protein HI0902

Query Match 46.7%; Score 49; DB 2; Length 251;
 Best Local Similarity 41.2%; Pred. No. 8.7;
 Matches 7; Conservative 6; Mismatches 4; Indels 0; Gaps 0;

Oy 4 VFTGVLAGVWGEGKEV 20
 :||:|||||
 Db 146 LFTGISIVAGAGAFI 162

RESULT 8

H69217
 hypothetical protein MTH882 - Methanobacterium thermoautotrophicum (strain Delta H)
 C:Species: Methanobacterium thermoautotrophicum
 C>Date: 05-Dec-1997 #sequence_revision 05-Dec-1997 #text_change 24-Sep-1999

C:Accession: H69217
 R:Smith, D.R.; Doucette-Stamm, L.A.; Deloughery, C.; Lee, H.; Dubois, J.; Aldredge, T.;
 R:Smith, D.; Spadafora, R.; Vicalote, R.; Wang, Y.; Wierzbowski, J.; Gibson, R.; Jiwani, N.
 K.; Church, G.M.; Daniels, C.J.; Mao, J.; Rice, P.; Noelling, J.; Reeve, J.N.
 J. Bacteriol. 179, 7135-7155, 1997

A>Title: Complete genome sequence of Methanobacterium thermoautotrophicum Delta H: func
 A:Reference number: A89000; MUID:96037514; PMID:9371463

A:Accession: H69217
 A>Status: preliminary; nucleic acid sequence not shown; translation not shown

A:Molecule type: DNA

A:Residues: 1-261 <MTH>

A:Cross-references: GB:AE000864; GB:AE000666; NID:92621970; PIDN:AA85380.1; PID:9262197

A:Experimental source: strain Delta H

A:Gene: MTH882

A:Start codon: TTG

C:Superfamily: hypothetical protein HI0902

Query Match 44.8%; Score 47; DB 2; Length 261;
 Best Local Similarity 50.0%; Pred. No. 18;
 Matches 11; Conservative 4; Mismatches 3; Indels 4; Gaps 1;

Oy 3 LVFTGVLA---GVMGEGKEV 20
 :||:|||||
 Db 12 LFTGLAGLATGILGVGGFI 33

RESULT 9

C69374
 conserved hypothetical protein AF0995 - Archaeoglobus fulgidus
 C:Species: Archaeoglobus fulgidus
 C>Date: 05-Dec-1997 #sequence_revision 05-Dec-1997 #text_change 22-Oct-1999

C:Accession: C69374
 R:Klenk, H.P.; Clayton, R.A.; Tomb, J.F.; White, O.; Nelson, K.F.; Ketchum, K.A.; Dodson
 R:Klenk, H.P.; Zhou, L.; Overbeek, R.; Gocayne, J.D.; Weidman, J.F.; McDonald, L.

Nature 390, 364-370, 1997
 A:Authors: Overbeek, T.; Cotton, M.D.; Spriggs, T.; Artlich, P.; Kaine, B.P.; Sykes, S.
 Smith, H.O.; Woese, C.R.; Venter, J.C.

A>Title: The complete genome sequence of the hyperthermophilic, sulfate-reducing archae

A:Reference number: A69250; MUID:98049343; PMID:9389475

A:Accession: C69374

A>Status: preliminary; nucleic acid sequence not shown; translation not shown

A:Molecule type: DNA

A:Residues: 1-651 <KLB>

A:Cross-references: GB:AE001035; GB:AE007082; NID:92689358; PIDN:AA80246.1; PID:9264960

Query Match 44.8%; Score 47; DB 2; Length 651;
 Best Local Similarity 61.5%; Pred. No. 43;
 Matches 8; Conservative 3; Mismatches 2; Indels 0; Gaps 0;

Oy 4 VFTGVLAGVWGEG 16

Db 610 IFAGMVGWVGEG 622

RESULT 10

G82446
 conserved hypothetical protein VCA0543 [imported] - Vibrio cholerae (strain N16961 serog

C:Species: Vibrio cholerae
 C>Date: 18-Aug-2000 #sequence_revision 20-Aug-2000 #text_change 02-Feb-2001

C:Accession: G82446
 R:Heidelberg, J.F.; Eisen, J.A.; Nelson, W.C.; Clayton, R.A.; Gwin, M.L.; Dodson, R.J.;
 chardson, D.; Ermolaeva, M.D.; Vamathevan, J.; Bass, S.; Qiu, H.; Dragol, I.; Sellers, P.
 L.; R.R.; Mekalanos, J.J.; Venter, J.C.; Fraser, C.M.

Nature 406, 477-483, 2000

A>Title: DNA sequence of both chromosomes of the cholera pathogen Vibrio cholerae.

A:Reference number: A82035; MUID:20406833; PMID:10952301

A:Accession: G82446

A>Status: preliminary

A:Molecule type: DNA

A:Residues: 1-1170 <HEI>

A:Cross-references: GB:AE004385; GB:AE003853; NID:9657946; PIDN:AA96445.1; GSPDB:GN001;
 A:Experimental source: serogroup O1; strain N16961; biotype El Tor

A:Gene: VCA0543

A:Map position: 2

Query Match 43.8%; Score 46; DB 2; Length 170;
 Best Local Similarity 77.8%; Pred. No. 17;
 Matches 7; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

Oy 6 TGVLAGVWG 14
 :||:|||||
 Db 16 TGLSGVWG 24

RESULT 11

AC0635
 probable membrane protein STY1175 [imported] - Salmonella enterica subsp. enterica serov

C:Species: Salmonella enterica subsp. enterica serovar Typh

A>Note: this species has also been called Salmonella typh

C>Date: 09-Nov-2001 #sequence_revision 09-Nov-2001 #text_change 18-Nov-2002

C:Accession: AC0635

R:Parkhill, J.; Dougan, G.; James, K.D.; Thomson, N.R.; Pickard, D.; Wain, J.; Churcher,
 th, T.; Connerton, P.; Cronin, A.; Davis, P.; Davies, R.M.; Dowd, L.; White, N.; Farrar,
 S.; Moule, S.; O'Garra, P.

Nature 413, 848-852, 2001
 A:Authors: Parry, C.; Quail, M.; Rutherford, K.; Simmonds, M.; Skelton, J.; Stevens, K.;
 A>Title: Complete genome sequence of a multiple drug resistant Salmonella enterica serov

A:Reference number: AB0502; MUID:21534947; PMID:11677608

A:Accession: AC0635

A>Status: preliminary

A:Molecule type: DNA

A:Residues: 1-176 <PAR>

A:Cross-references: GB:AL513382; PIDN:CAD08262.1; PID:916502309; GSPDB:GN00476

A:Gene: STY1175

Query Match 43.8%; Score 46; DB 2; Length 176;
 Best Local Similarity 54.5%; Pred. No. 17;
 Matches 6; Conservative 4; Mismatches 1; Indels 0; Gaps 0;

Oy 4 VFTGVLAGVWG 14
 :||:|||||
 Db 25 ITTGLSGVWG 35

RESULT 12

A64846
 probable membrane protein ycdZ - Escherichia coli (strain K-12)

C:Species: Escherichia coli
 C>Date: 12-Sep-1997 #sequence_revision 17-Sep-1997 #text_change 01-Mar-2002

C:Accession: A64846

R:Blattner, F.R.; Plunkett III, G.; Bloch, C.A.; Perna, N.T.; Burland, V.; Riley, M.; Co

.A.: Rose, D.J.; Mau, B.; Shao, Y.
 Science 277, 1453-1462, 1997
 A>Title: The complete genome sequence of *Escherichia coli* K-12.
 A:Reference number: A64720; MUID:97426617; PMID:9278503
 A:Accession: A64846
 A>Status: nucleic acid sequence not shown; translation not shown
 A:Molecule type: DNA
 A:Residues: 1-179 <BLAT>
 A:Cross-references: GB:AB000205; GB:U00096; NID:G1787265; PIDN:AAC74120.1; PID:G1787273;
 A:Experimental source: strain K-12, substrain MG1655
 C:Genetics:
 A:Gene: ycdZ
 C:Keywords: transmembrane protein
 F:19-35/Domain: transmembrane #status predicted <TM1>
 F:37-53/Domain: transmembrane #status predicted <TM2>
 F:68-84/Domain: transmembrane #status predicted <TM3>
 F:94-110/Domain: transmembrane #status predicted <TM4>
 F:115-131/Domain: transmembrane #status predicted <TM5>
 F:138-155/Domain: transmembrane #status predicted <TM6>

Query Match 43.8%; Score 46; DB 2; Length 179;
 Best Local Similarity 54.5%; Pred. No. 17;
 Matches 6; Conservative 4; Mismatches 1; Indels 0; Gaps 0;

QY 4 VFTGVLAVGWG 14
 : |||: |||
 Db 25 ITTGILSGIWG 35

RESULT 13
 E90805
 Hypothetical protein Ecs1413 [imported] - *Escherichia coli* (strain O157:H7, substrain R1
 C:Species: *Escherichia coli*
 C>Date: 18-Jul-2001 #sequence_revision 18-Jul-2001 #text_change 18-Jul-2001
 C:Accession: E90805
 R:Hayashi, T.; Makino, K.; Ohnishi, M.; Kurokawa, K.; Ishii, K.; Yokoyama, K.; Han, C.G.
 Gasawara, N.; Yasunaga, T.; Kuhara, S.; Shiba, T.; Hatori, M.; Shinagawa, H.
 DNA Res. 8, 11-22, 2001
 A>Title: Complete genome sequence of enterohemorrhagic *Escherichia coli* O157:H7 and genc
 A:Reference number: A96629; MUID:21156231; PMID:11258796
 A:Accession: E90805
 A>Status: Preliminary
 A:Molecule type: DNA
 A:Residues: 1-179 <HAY>
 A:Cross-references: GB:BA000007; PIDN:BA34836.1; PID:G13360873; GSPDB:GN00154
 A:Experimental source: strain O157:H7, substrain R1MD 05095952
 C:Genetics:
 A:Gene: Ecs1413

Query Match 43.8%; Score 46; DB 2; Length 179;
 Best Local Similarity 54.5%; Pred. No. 17;
 Matches 6; Conservative 4; Mismatches 1; Indels 0; Gaps 0;

QY 4 VFTGVLAVGWG 14
 : |||: |||
 Db 25 ITTGILSGIWG 35

RESULT 14
 B85665
 Hypothetical protein ycdZ [imported] - *Escherichia coli* (strain O157:H7, substrain EDL93
 C:Species: *Escherichia coli*
 C>Date: 16-Feb-2001 #sequence_revision 16-Feb-2001 #text_change 14-Sep-2001
 C:Accession: B85665
 R:Perna, N.T.; Plunkett III, G.; Burland, V.; Mau, B.; Glasner, J.D.; Rose, D.J.; Mayhew
 Miller, L.; Grobeck, E.J.; Davis, N.W.; Iam, A.; Dimalanta, E.; Potamousis, K.; Apodaca,
 Nature 409, 529-533, 2001
 A>Title: Genome sequence of enterohemorrhagic *Escherichia coli* O157:H7.
 A:Reference number: A85480; MUID:21074935; PMID:11206551
 A:Accession: B85665
 A>Status: Preliminary
 A:Molecule type: DNA
 A:Residues: 1-179 <STO>

A:Cross-references: GB:AB005174; NID:G12514565; PIDN:AAG55782.1; GSPDB:GN00145; UMGF:Z166
 A:Experimental source: strain O157:H7, substrain EDL933
 C:Genetics:
 A:Gene: ycdZ

Query Match 43.8%; Score 46; DB 2; Length 179;
 Best Local Similarity 54.5%; Pred. No. 17;
 Matches 6; Conservative 4; Mismatches 1; Indels 0; Gaps 0;

QY 4 VFTGVLAVGWG 14
 : |||: |||
 Db 25 ITTGILSGIWG 35

RESULT 15
 AF0681
 Probable voltage gated chloride channel protein STY1574 [imported] - *Salmonella enterica*
 C:Species: *Salmonella enterica* subsp. *enterica* serovar Typh
 A>Note: this species has also been called *Salmonella typhi*
 C>Date: 09-Nov-2001 #sequence_revision 09-Nov-2001 #text_change 18-Nov-2002
 C:Accession: AF0681
 R:Parkhill, J.; Dougan, G.; James, K.D.; Thomson, N.R.; Pickard, D.; Wain, J.; Churcher,
 Th, T.; Connor, P.; Cronin, A.; Davis, P.; Davies, R.M.; Dowd, L.; White, N.; Farrar,
 S.; Moule, S.; O'Gaora, P.
 Nature 413, 848-852, 2001
 A:Authors: Parry, C.; Quail, M.; Rutherford, K.; Simmonds, M.; Skelton, J.; Stevens, K.;
 A>Title: Complete genome sequence of a multiple drug resistant *Salmonella enterica* serove
 A:Reference number: AB0502; MUID:21534947; PMID:11677608
 A:Accession: AF0681
 A>Status: Preliminary
 A:Molecule type: DNA
 A:Residues: 1-429 <PAR>
 A:Cross-references: GB:AL513382; PIDN:CAD01823.1; PID:G16502667; GSPDB:GN00176
 C:Genetics:
 A:Gene: STY1574
 C:Superfamily: hypothetical protein s110855

Query Match 43.8%; Score 46; DB 2; Length 429;
 Best Local Similarity 50.0%; Pred. No. 41;
 Matches 7; Conservative 5; Mismatches 2; Indels 0; Gaps 0;

QY 1 TELVFTGVLAVGWG 14
 : |||: |||
 Db 234 VMIIVSTGVLAVGLWG 247

Search completed: September 30, 2004, 11:09:47
 Job time : 14.7619 secs

CC viral component. The antiviral agent can be used to combat herpes virus
 CC replication or infection

SQ Sequence 20 AA;

Query Match 100.0%; Score 105; DB 2; Length 20;
 Best Local Similarity 100.0%; Pred. No. 4.1e-09;
 Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 IELVFTGLAGVWEGGKRV 20
 |||||
 DB 1 IELVFTGLAGVWEGGKRV 20

RESULT 2

AAW53265
 ID AAW53265 standard; peptide; 29 AA.

AC AAW53265;

DT 01-JUL-1998 (first entry)

DE Herpes simplex virus type 1 antiviral agent peptide 5.

KM Herpes simplex virus type 1; HSV-1; antiviral; inhibition; infection;
 KM replication; UL8; POL.

OS Synthetic.

OS Herpes simplex virus unknown type.

PN W09804707-A1.

PD 05-FEB-1998.

PF 28-JUL-1997; 97WO-GB002025.

PR 26-JUL-1996; 96GB-00015730.

PA (MED1-) MEDICAL RES COUNCIL.

PI Marsden HS, Stow ND, McLean GW;

DR WPI, 1998-130695/12.

PT Antiviral agent capable of inhibiting herpes virus replication - disrupts
 association between herpes simplex virus type 1 UL8 and POL.

PS Claim 5; Page 57; 83pp; English.

CC The present sequence represents an antiviral agent peptide capable of
 combating herpes virus replication. The antiviral agent is capable of
 disrupting the association between UL8 and POL (UL30), where UL8 and POL
 are respectively defined as UL8 and POL of herpes simplex virus type 1
 (HSV-1) together with homologues in other herpes viruses. The present
 invention also describes an assay to determine the ability of a test
 substance to interfere with the association of UL8 and POL. The assay
 comprises: (a) exposing a 1st viral component to a test substance
 followed by a 2nd viral component, or exposing a 1st viral component to a
 2nd viral component followed by a test substance; (b) washing to remove
 any 2nd viral component and/or test substance not associated with the 1st
 viral component; and (c) detecting the presence, and optionally
 determining the amount, of 2nd viral component associated with the 1st
 viral component. The antiviral agent can be used to combat herpes virus
 replication or infection

SQ Sequence 29 AA;

Query Match 87.6%; Score 92; DB 2; Length 29;
 Best Local Similarity 100.0%; Pred. No. 4.3e-06;
 Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 4 VFTGLAGVWEGGKRV 20
 |||||

DB 1 VFTGLAGVWEGGKRV 17

RESULT 3

AAW72064
 ID AAW72064 standard; protein; 386 AA.

AC AAW72064;

DT 18-DEC-1998 (first entry)

DE HSV-2 strain SBS Contig ID 93 ORF#2 protein.

KM HSV-2 strain SBS; immunological response induction; therapy;
 KM antiviral identification; viral protein inhibitor.

OS Herpes simplex virus 2.

PN W09820016-A1.

PD 14-MAY-1998.

PF 31-OCT-1997; 97WO-US020016.

PR 04-NOV-1996; 96US-0030279P.

PR 09-JUN-1997; 97US-0049018P.

PA (SMIK) SMITHKLINE BEECHAM CORP.

PI Esser KM, Chan JY, Dabrowski Amaral CE, Delvecchio AM, Dillon SB;

PI Leary JG;

DR WPI; 1998-286847/25.

DR N-PSDB; AAW62147.

PT Herpes simplex virus type-2 sequences - useful in, e.g. prevention and
 treatment of infection or inducing immunological response in mammal.

PS Claim 10; Page 65; 748pp; English.

CC This sequence represents a Herpes simplex virus type-2 (HSV-2) protein
 CC sequence of the invention. This sequence was isolated from a HSV-2 strain
 CC SBS (deposited as ATCC VR-2546) DNA fragment designated Contig ID 93.
 CC Based on homology, this sequence is a UL8 protein. The proteins can be
 CC used for the treatment or prevention of disease, to induce an
 CC immunological response in a mammal or to identify inhibitors, activators
 CC or novel antivirals. Antagonists of the proteins can be used to inhibit a
 CC viral polypeptide. The DNA sequence or a vector containing it can also be
 CC used to induce an immunological response in a mammal

SQ Sequence 386 AA;

Query Match 87.6%; Score 92; DB 2; Length 386;
 Best Local Similarity 85.0%; Pred. No. 5.8e-05;
 Matches 17; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 1 IELVFTGLAGVWEGGKRV 20
 |||||
 DB 355 IQLVFTGLAGVWEGGKRV 374

RESULT 4

AAW72138
 ID AAW72138 standard; protein; 408 AA.

AC AAW72138;

DT 23-DEC-1998 (first entry)

DE HSV-2 strain SBS Contig ID 18 ORF#3 protein.

KM HSV-2 strain SBS; immunological response induction; therapy;
 KM antiviral identification; viral protein inhibitor.


```

XX XX      Herpes simplex virus 2.
OS
XX
XX Key      Location/Qualifiers
FH Misc-difference 1
FT /note= "encoded by CG"
XX
XX WO9820016-A1.
XX
XX 14-MAY-1998.
XX
XX 31-OCT-1997; 97WO-US020016.
XX
XX 04-NOV-1996; 96US-0030279P.
XX
XX 09-JUN-1997; 97US-0049018P.
XX
XX (SMIK ) SMITHKLINE BEECHAM CORP.
XX
XX Esser KM, Chan JY, Dabrowski-Amara CE, Delvecchio AM, Dillon SB,
XX Leary JY;
XX
XX WP1; 1998-286847/25.
XX
XX N-PSDB; AAV62162.
XX
XX Herpes simplex virus type-2 sequences - useful in, e.g. prevention and
XX treatment of infection or inducing immunological response in mammal.
XX
XX Claim 10; Page 97; 748pp; English.
XX
XX This sequence represents a Herpes simplex virus type-2 (HSV-2) protein
XX sequence of the invention. This sequence was isolated from a HSV-2 strain
XX SB5 (deposited as ATCC VR-2546) DNA fragment designated Contig ID 18. The
XX proteins can be used for the treatment or prevention of disease, to
XX induce an immunological response in a mammal or to identify inhibitors,
XX activators or novel antivirals. Antagonists of the proteins can be used
XX to inhibit a viral polypeptide. The DNA sequence or a vector containing
XX it can also be used to induce an immunological response in a mammal
XX
XX Sequence 408 AA;
SQ
Query Match      87.6%; Score 92; DB 2; Length 408;
Best Local Similarity 85.0%; Pred. No. 6.1e-05;
Matches 17; Conservative 1; Mismatches 2; Indels 0; Gaps 0;
OY      1 IEIVFTGVLGVWGGGKFEV 20
Db      377 IQLVFTGVLEGVWGGGSEFV 396

```

```

XX XX      04-NOV-1996; 96US-0030279P.
XX PR 09-JUN-1997; 97US-0049018P.
XX
XX (SMIK ) SMITHKLINE BEECHAM CORP.
XX
XX Esser KM, Chan JY, Dabrowski-Amara CE, Delvecchio AM, Dillon SB,
XX Leary JY;
XX
XX WP1; 1998-286847/25.
XX
XX N-PSDB; AAV62162.
XX
XX Herpes simplex virus type-2 sequences - useful in, e.g. prevention and
XX treatment of infection or inducing immunological response in mammal.
XX
XX Claim 10; Page 112; 748pp; English.
XX
XX This sequence represents a Herpes simplex virus type-2 (HSV-2) protein
XX sequence of the invention. This sequence was isolated from a HSV-2 strain
XX SB5 (deposited as ATCC VR-2546) DNA fragment designated Contig ID 15. The
XX proteins can be used for the treatment or prevention of disease, to
XX induce an immunological response in a mammal or to identify inhibitors,
XX activators or novel antivirals. Antagonists of the proteins can be used
XX to inhibit a viral polypeptide. The DNA sequence or a vector containing
XX it can also be used to induce an immunological response in a mammal
XX
XX Sequence 750 AA;
SQ
Query Match      87.6%; Score 92; DB 2; Length 750;
Best Local Similarity 85.0%; Pred. No. 0.00011;
Matches 17; Conservative 1; Mismatches 2; Indels 0; Gaps 0;
OY      1 IEIVFTGVLGVWGGGKFEV 20
Db      719 IQLVFTGVLEGVWGGGSEFV 738

```

RESULT 6

```

AAW53273
ID AAW53273 standard; peptide; 15 AA.
XX
XX AAW53273;
XX
XX 01-JUL-1998 (first entry)
XX
XX Herpes simplex virus type 1 antiviral peptide 6.
XX
XX Herpes simplex virus type 1; HSV-1; antiviral; inhibition; infection;
XX replication; UL8; POL.
XX
XX Synthetic.
XX
XX Herpes simplex virus unknown type.
XX
XX WO9804707-A1.
XX
XX 05-FEB-1998.
XX
XX 28-JUL-1997; 97WO-GB002025.
XX
XX 26-JUL-1996; 96GB-00015730.
XX
XX (MEDI-) MEDICAL RES COUNCIL.
XX
XX Marsden HS, Stow ND, McLean GW;
XX
XX WP1; 1998-130695/12.
XX
XX Antiviral agent capable of inhibiting herpes virus replication - disrupts
XX association between herpes simplex virus type 1 UL8 and POL.
XX
XX Example 1; Page 43; 83pp; English.
XX
XX The present sequence represents a peptide used in an example of the
XX

```

CC present invention. The present invention describes an antiviral agent
 CC capable of combatting herpes virus replication. The antiviral agent is
 CC capable of disrupting the association between UL8 and POL (UL30), where
 CC UL8 and POL are respectively defined as UL8 and POL of herpes simplex
 CC virus type 1 (HSV-1) together with homologues in other herpes viruses.
 CC The present invention also describes an assay to determine the ability of
 CC a test substance to interfere with the association of UL8 and POL. The
 CC assay comprises: (a) exposing a 1st viral component to a test substance
 CC followed by a 2nd viral component, or exposing a 1st viral component to a
 CC 2nd viral component followed by a test substance; (b) washing to remove
 CC any 2nd viral component and/or test substance not associated with the 1st
 CC viral component; and (c) detecting the presence, and optionally
 CC determining the amount, of 2nd viral component associated with the 1st
 CC viral component. The antiviral agent can be used to combat herpes virus
 CC replication or infection

SQ Sequence 15 AA;

Query Match 78.1%; Score 82; DB 2; Length 15;
 Best Local Similarity 100.0%; Pred. No. 6e-05;
 Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 6 TGVLAGVWGEGRFV 20
 DB 1 TGVLAGVWGEGRFV 15

RESULT 7

AAW53272
 ID AAW53272 standard; peptide; 27 AA.

AAW53272;

01-UTL-1998 (first entry)

DE Herpes simplex virus type 1 antiviral peptide 4.

KW Herpes simplex virus type 1; HSV-1; antiviral; inhibition; infection;

KM replication; UL8; POL.

OS Synthetic.

OS Herpes simplex virus unknown type.

PN WO9804707-A1.

PD 05-FEB-1998.

PE 28-UTL-1997; 97WO-GB002025.

PR 26-JUL-1996; 96GB-00015730.

PS (MED1-) MEDICAL RES COUNCIL.

PI Marsden HS, Stow ND, Mclean GW;

PT WPI; 1998-130695/12.

DR Antiviral agent capable of inhibiting herpes virus replication - disrupts

PT association between herpes simplex virus type 1 UL8 and POL.

XX Example 1; Page 43; 83pp; English.

XX The present sequence represents a peptide used in an example of the
 CC present invention. The present invention describes an antiviral agent
 CC capable of combatting herpes virus replication. The antiviral agent is
 CC capable of disrupting the association between UL8 and POL (UL30), where
 CC UL8 and POL are respectively defined as UL8 and POL of herpes simplex
 CC virus type 1 (HSV-1) together with homologues in other herpes viruses.
 CC The present invention also describes an assay to determine the ability of
 CC a test substance to interfere with the association of UL8 and POL. The
 CC assay comprises: (a) exposing a 1st viral component to a test substance
 CC followed by a 2nd viral component, or exposing a 1st viral component to a
 CC 2nd viral component followed by a test substance; (b) washing to remove

CC any 2nd viral component and/or test substance not associated with the 1st
 CC viral component; and (c) detecting the presence, and optionally
 CC determining the amount, of 2nd viral component associated with the 1st
 CC viral component. The antiviral agent can be used to combat herpes virus
 CC replication or infection

SQ Sequence 27 AA;

Query Match 78.1%; Score 82; DB 2; Length 27;
 Best Local Similarity 100.0%; Pred. No. 0.00011;
 Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 6 TGVLAGVWGEGRFV 20
 DB 1 TGVLAGVWGEGRFV 15

RESULT 8

AAW53271
 ID AAW53271 standard; peptide; 25 AA.

AAW53271;

01-UTL-1998 (first entry)

DE Herpes simplex virus type 1 antiviral peptide 3.

KW Herpes simplex virus type 1; HSV-1; antiviral; inhibition; infection;

KM replication; UL8; POL.

OS Synthetic.

OS Herpes simplex virus unknown type.

PN WO9804707-A1.

PD 05-FEB-1998.

PE 28-UTL-1997; 97WO-GB002025.

PR 26-JUL-1996; 96GB-00015730.

PS (MED1-) MEDICAL RES COUNCIL.

PI Marsden HS, Stow ND, Mclean GW;

PT WPI; 1998-130695/12.

DR Antiviral agent capable of inhibiting herpes virus replication - disrupts

PT association between herpes simplex virus type 1 UL8 and POL.

XX Example 1; Page 43; 83pp; English.

XX The present sequence represents a peptide used in an example of the
 CC present invention. The present invention describes an antiviral agent
 CC capable of combatting herpes virus replication. The antiviral agent is
 CC capable of disrupting the association between UL8 and POL (UL30), where
 CC UL8 and POL are respectively defined as UL8 and POL of herpes simplex
 CC virus type 1 (HSV-1) together with homologues in other herpes viruses.
 CC The present invention also describes an assay to determine the ability of
 CC a test substance to interfere with the association of UL8 and POL. The
 CC assay comprises: (a) exposing a 1st viral component to a test substance
 CC followed by a 2nd viral component, or exposing a 1st viral component to a
 CC 2nd viral component followed by a test substance; (b) washing to remove
 CC any 2nd viral component and/or test substance not associated with the 1st
 CC viral component; and (c) detecting the presence, and optionally
 CC determining the amount, of 2nd viral component associated with the 1st
 CC viral component. The antiviral agent can be used to combat herpes virus
 CC replication or infection

SQ Sequence 25 AA;

Query Match 67.6%; Score 71; DB 2; Length 25;
 Best Local Similarity 100.0%; Pred. No. 0.0038;

Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 8 VLAGVMBGGKRV 20
| | | | | | | | | |
Db 1 VLAGVMBGGKRV 13

RESULT 9

AAW53270 ID AAW53270 standard; peptide: 23 AA.

AAW53270;

01-JUL-1998 (first entry)

Herpes simplex virus type 1 antiviral peptide 2.

Herpes simplex virus type 1; HSV-1; antiviral; inhibition; infection;

replication; UL8; POL.

Synthetic.

Herpes simplex virus unknown type.

WO9804707-A1.

05-FEB-1998.

28-JUL-1997; 97WO-GB002025.

26-JUL-1996; 96GB-00015730.

(MED1-) MEDICAL RES COUNCIL.

Mareiden HS, Stow ND, Mclean GW;

WPI; 1998-130695/12.

Antiviral agent capable of inhibiting herpes virus replication - disrupts association between herpes simplex virus type 1 UL8 and POL.

Example 1; Page 43; 83pp; English.

The present sequence represents a peptide used in an example of the present invention. The present invention describes an antiviral agent capable of combating herpes virus replication. The antiviral agent is capable of disrupting the association between UL8 and POL (UL30), where UL8 and POL are respectively defined as UL8 and POL of herpes simplex virus type 1 (HSV-1) together with homologues in other herpes viruses. The present invention also describes an assay to determine the ability of a test substance to interfere with the association of UL8 and POL. The assay comprises: (a) exposing a 1st viral component to a test substance followed by a 2nd viral component, or exposing a 1st viral component to a 2nd viral component followed by a test substance; (b) washing to remove any 2nd viral component and/or test substance not associated with the 1st viral component; and (c) detecting the presence, and optionally determining the amount, of 2nd viral component associated with the 1st viral component. The antiviral agent can be used to combat herpes virus replication or infection

Sequence 23 AA;

Query Match 60.0%; Score 63; DB 2; Length 23;

Best Local Similarity 100.0%; Pred. No. 0.049; Mismatches 0; Indels 0; Gaps 0;

QY 10 AGVMBGGKRV 20
| | | | | | | | | |
Db 1 AGVMBGGKRV 11

RESULT 10

AAW71478 ID AAW71478 standard; protein: 255 AA.

AAW71478;

09-NOV-1998 (first entry)

Helicobacter polypeptide GHPO 136.

GHPO 136; infection; therapy; diagnosis; vaccine; gastritis; ulcer.

Helicobacter pylori.

WO9821225-A1.

22-MAY-1998.

14-NOV-1997; 97WO-US021353.

14-NOV-1996; 96US-00749051.

01-APR-1997; 97US-00831309.

01-APR-1997; 97US-00834457.

01-APR-1997; 97US-00834705.

24-JUN-1997; 97US-00881227.

29-JUL-1997; 97US-00902615.

(INMR) MERIEUX ORAVAX PASTEUR MERIEUX SERUMS.

(PLAC) MAX PLANCK GES FORDERUNG WISSENSCHAFTEN.

(HUMA-) HUMAN GENOME SCI INC.

Haas R, Kleantous H, Tomb J, Miller C, Al-Garawi A, Odenbreit S; Meyer T;

WPI; 1998-297855/26.

Helicobacter polynucleotide and polypeptide sequences - useful to treat or prevent gastrointestinal infection.

Claim 1; Page 90-91; 362pp; English.

This claimed Helicobacter pylori polypeptide, designated GHPO 136, can be used in vaccination methods for preventing or treating Helicobacter infection. 85 Helicobacter polypeptides (see AAW71474-W71558) are claimed, as well as isolated polynucleotides (see AAV52009-93) that encode them. The invention also provides: methods for producing these expression cassettes, vectors and recombinant host systems, and related live vaccine vectors that contain the polynucleotides of the invention and which can be used to prevent or treat Helicobacter infection; CC therapeutic and/or prophylactic methods involving administration of polynucleotide molecules, polypeptides or monospecific antibodies; CC methods for detecting the presence of Helicobacter in samples using e.g. the polypeptides or monospecific antibodies; and methods for purifying the polypeptides by antibody-based affinity chromatography

Sequence 255 AA;

Query Match 49.5%; Score 52; DB 2; Length 255;

Best Local Similarity 47.1%; Pred. No. 20; Mismatches 3; Indels 0; Gaps 0;

QY 4 VFTGVLAGVMBGGKRV 20
| | | | | | | | | |
Db 12 LFTGVLAGVMBGGKRV 18

RESULT 11

AAW46309 ID AAW46309 standard; protein: 255 AA.

AAW46309;

05-APR-2001 (first entry)


```
XX WO200157182-A2.
PN
XX
PD 09-AUG-2001.
PF 17-JAN-2001; 2001WC-US001354.
XX
XX 31-JAN-2000; 2000US-0179065P.
PR 04-FEB-2000; 2000US-0180628P.
PR 24-FEB-2000; 2000US-0184664P.
PR 02-MAR-2000; 2000US-0186350P.
PR 16-MAR-2000; 2000US-0189874P.
PR 17-MAR-2000; 2000US-0190076P.
PR 18-APR-2000; 2000US-0198123P.
PR 19-MAY-2000; 2000US-0205515P.
PR 07-JUN-2000; 2000US-0209467P.
PR 28-JUN-2000; 2000US-0214886P.
PR 30-JUN-2000; 2000US-0215135P.
PR 07-JUL-2000; 2000US-0216647P.
PR 07-JUL-2000; 2000US-0216880P.
PR 11-JUL-2000; 2000US-0217487P.
PR 11-JUL-2000; 2000US-0217496P.
PR 14-JUL-2000; 2000US-0218290P.
PR 26-JUL-2000; 2000US-0220963P.
PR 26-JUL-2000; 2000US-0220964P.
PR 14-AUG-2000; 2000US-0224518P.
PR 14-AUG-2000; 2000US-0224519P.
PR 14-AUG-2000; 2000US-0225213P.
PR 14-AUG-2000; 2000US-0225214P.
PR 14-AUG-2000; 2000US-0225266P.
PR 14-AUG-2000; 2000US-0225267P.
PR 14-AUG-2000; 2000US-0225268P.
PR 14-AUG-2000; 2000US-0225272P.
PR 14-AUG-2000; 2000US-0225347P.
PR 14-AUG-2000; 2000US-0225757P.
PR 14-AUG-2000; 2000US-0225756P.
PR 14-AUG-2000; 2000US-0225759P.
PR 18-AUG-2000; 2000US-0226279P.
PR 22-AUG-2000; 2000US-0226681P.
PR 22-AUG-2000; 2000US-0226686P.
PR 22-AUG-2000; 2000US-0227182P.
PR 23-AUG-2000; 2000US-0227009P.
PR 30-AUG-2000; 2000US-0228924P.
PR 01-SEP-2000; 2000US-0229287P.
PR 01-SEP-2000; 2000US-0229343P.
PR 01-SEP-2000; 2000US-0229344P.
PR 01-SEP-2000; 2000US-0229345P.
PR 05-SEP-2000; 2000US-0229509P.
PR 05-SEP-2000; 2000US-0229513P.
PR 06-SEP-2000; 2000US-0230437P.
PR 08-SEP-2000; 2000US-0230438P.
PR 08-SEP-2000; 2000US-0231243P.
PR 08-SEP-2000; 2000US-0231244P.
PR 08-SEP-2000; 2000US-0231413P.
PR 08-SEP-2000; 2000US-0231414P.
PR 08-SEP-2000; 2000US-0232080P.
PR 08-SEP-2000; 2000US-0232081P.
PR 12-SEP-2000; 2000US-0231968P.
PR 14-SEP-2000; 2000US-0233397P.
PR 14-SEP-2000; 2000US-0233398P.
PR 14-SEP-2000; 2000US-0233399P.
PR 14-SEP-2000; 2000US-0233400P.
PR 14-SEP-2000; 2000US-0233401P.
PR 14-SEP-2000; 2000US-0233663P.
PR 14-SEP-2000; 2000US-0233664P.
PR 14-SEP-2000; 2000US-0233665P.
PR 21-SEP-2000; 2000US-0234223P.
PR 21-SEP-2000; 2000US-0234274P.
PR 25-SEP-2000; 2000US-0234997P.
PR 25-SEP-2000; 2000US-0234998P.
PR 26-SEP-2000; 2000US-0235484P.
PR 27-SEP-2000; 2000US-0235834P.
XX
XX 27-SEP-2000; 2000US-0235836P.
PR 29-SEP-2000; 2000US-0236327P.
PR 29-SEP-2000; 2000US-0236367P.
PR 29-SEP-2000; 2000US-0236368P.
PR 29-SEP-2000; 2000US-0236369P.
PR 29-SEP-2000; 2000US-0236370P.
PR 02-OCT-2000; 2000US-0236802P.
PR 02-OCT-2000; 2000US-0237037P.
PR 02-OCT-2000; 2000US-0237038P.
PR 02-OCT-2000; 2000US-0237039P.
PR 02-OCT-2000; 2000US-0237040P.
PR 13-OCT-2000; 2000US-0239335P.
PR 13-OCT-2000; 2000US-0239337P.
PR 20-OCT-2000; 2000US-0240960P.
PR 20-OCT-2000; 2000US-0241221P.
PR 20-OCT-2000; 2000US-0241785P.
PR 20-OCT-2000; 2000US-0241786P.
PR 20-OCT-2000; 2000US-0241787P.
PR 20-OCT-2000; 2000US-0241808P.
PR 20-OCT-2000; 2000US-0241809P.
PR 20-OCT-2000; 2000US-0241826P.
PR 01-NOV-2000; 2000US-0244617P.
PR 08-NOV-2000; 2000US-0246474P.
PR 08-NOV-2000; 2000US-0246475P.
PR 08-NOV-2000; 2000US-0246476P.
PR 08-NOV-2000; 2000US-0246477P.
PR 08-NOV-2000; 2000US-0246478P.
PR 08-NOV-2000; 2000US-0246523P.
PR 08-NOV-2000; 2000US-0246524P.
PR 08-NOV-2000; 2000US-0246525P.
PR 08-NOV-2000; 2000US-0246526P.
PR 08-NOV-2000; 2000US-0246527P.
PR 08-NOV-2000; 2000US-0246528P.
PR 08-NOV-2000; 2000US-0246532P.
PR 08-NOV-2000; 2000US-0246609P.
PR 08-NOV-2000; 2000US-0246610P.
PR 08-NOV-2000; 2000US-0246611P.
PR 08-NOV-2000; 2000US-0246613P.
PR 17-NOV-2000; 2000US-0249207P.
PR 17-NOV-2000; 2000US-0249208P.
PR 17-NOV-2000; 2000US-0249209P.
PR 17-NOV-2000; 2000US-0249210P.
PR 17-NOV-2000; 2000US-0249211P.
PR 17-NOV-2000; 2000US-0249212P.
PR 17-NOV-2000; 2000US-0249213P.
PR 17-NOV-2000; 2000US-0249214P.
PR 17-NOV-2000; 2000US-0249264P.
PR 17-NOV-2000; 2000US-0249265P.
PR 17-NOV-2000; 2000US-0249297P.
PR 17-NOV-2000; 2000US-0249299P.
PR 17-NOV-2000; 2000US-0249300P.
PR 01-DEC-2000; 2000US-0250160P.
PR 01-DEC-2000; 2000US-0250391P.
PR 05-DEC-2000; 2000US-0251030P.
PR 05-DEC-2000; 2000US-0251988P.
PR 05-DEC-2000; 2000US-0256719P.
PR 06-DEC-2000; 2000US-0251479P.
PR 08-DEC-2000; 2000US-0251868P.
PR 08-DEC-2000; 2000US-0251869P.
PR 08-DEC-2000; 2000US-0251989P.
PR 11-DEC-2000; 2000US-0254097P.
PR 05-JAN-2001; 2001US-0259678P.
XX
XX (HUMA-) HUMAN GENOME SCI INC.
XX
```



```

FT      /note= "ketoreductase domain"
FT      3627. .3710
FT      /label= ACP4 domain
FT      /note= "Acyl carrier protein domain"
XX      WO200026349-A2.
XX      11-MAY-2000.
XX      22-OCT-1999; 99WO-US024478.
XX      29-OCT-1998; 98US-0106100P.
XX      16-FEB-1999; 99US-0120254P.
XX      (KOSA-) KOSAN BIOSCIENCES INC.
XX      Belach MC, Shah SK, McDaniel R, Tang L;
XX      WPI: 2000-365602/31.
XX      N-PSDB; AAA09469.
XX      Recombinant DNA compound encoding oleandolide polyketide synthase for
PT      synthesizing polyketides comprising a coding sequence for a domain of a
PT      loading module or any one of extender modules.
XX      Disclosure: Page 28-29; 86pp; English.
XX      The oleandolide polyketide synthase (PKS), also known as 8,8a-
CC      deoxyoleandolide synthase, is encoded by three open reading frames (ORF),
CC      designated oleAI, oleAII and oleAIII. The PKS is a type I "modular"
CC      enzyme, where each ORF encodes 2 extender modules and the first ORF also
CC      encodes the loading module. Each module is composed of at least a
CC      ketosynthase (KS), acyl-transferase (AT) and an acyl carrier protein
CC      (ACP) domain. The oleandolide PKS loading module contains an inactivated
CC      KS, called KS-Q, where Q is the abbreviation for glutamine, present
CC      instead of the active site cysteine required for activity. The large
CC      multifunctional PKS enzymes catalyze the biosynthesis of polyketide
CC      macrolactones through multistep pathways involving decarboxylative
CC      condensations between acylthioesters followed by cycles of varying beta-
CC      carbon processing activities. The macrolide product of the PKS, 8,8a-
CC      deoxyoleandolide, is further modified by epoxidation and glycosylation to
CC      yield oleandomycin, an antibacterial polyketide. The invention concerns
CC      an isolated recombinant DNA compound, comprising a coding sequence for a
CC      domain of loading module or any one of extender modules 1-4 or 1-6,
CC      including an oleandolide PKS operably linked to a promoter. Also
CC      discussed are recombinant oleandolide PKS in which the module 1 KS domain
CC      is inactivated by deletion or other mutation. In particular, the
CC      inactivation is mediated by a change in the KS domain that renders it
CC      incapable of binding substrate (the KS1-Q mutation), rendered by mutation
CC      in the codon for the active site cysteine. The oleandolide PKS is useful
CC      for synthesizing polyketides, which are useful as antibiotics and
CC      molluscs. Heterologous expression of oleandolide PKS in host cells such
CC      as Streptomyces coelicolor and S. lividans is also made possible.
CC      Unmodified oleandolide compounds can be provided to cultures of
CC      Saccharopolyspora erythraea and converted to the corresponding
CC      derivatives of erythromycins A-D. (Updated on 06-AUG-2003 to correct OS
CC      field.)
XX      SQ      Sequence 3816 AA:
XX
XX      Query Match      45.7%; Score 48; DB 3; Length 3816;
XX      Best Local Similarity 47.1%; Pred. No. 1.1e+03;
XX      Matches 8; Conservative 4; Mismatches 5; Indels 0; Gaps 0;
XX      2 ELVFTGLAGVWGEGK 18
XX      :|||:|||||:
XX      1341 DVFFSSVAGVWGSAGQ 1357

```

Search completed: September 30, 2004, 11:04:47
 Job time : 53.8095 secs

This Page Blank (uspto)

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OK protein - protein search, using sw model

Run on: September 30, 2004, 11:08:46 ; Search time 72.9143 Seconds
(without alignments)
127.988 Million cell updates/sec

Title: US-09-988-851a-5
Sequence: 1 VFTGVLAGVWGEGKRVYPPDDKMSFLFA 29

Scoring table: BLOSUM62
Gapop 10.0, Gapext 0.5

Searched: 1351062 seqs, 32179191 residues

Total number of hits satisfying chosen parameters: 1351062

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database: Published Applications_AA:*

```
1: /cgn2_6/ptodata/1/pubpaa/US07_PUBCOMB.pep:*
2: /cgn2_6/ptodata/1/pubpaa/PCT_NEW_PUB.pep:*
3: /cgn2_6/ptodata/1/pubpaa/US06_NEW_PUB.pep:*
4: /cgn2_6/ptodata/1/pubpaa/US06_PUBCOMB.pep:*
5: /cgn2_6/ptodata/1/pubpaa/US07_NEW_PUB.pep:*
6: /cgn2_6/ptodata/1/pubpaa/PCTUS_PUBCOMB.pep:*
7: /cgn2_6/ptodata/1/pubpaa/US08_NEW_PUB.pep:*
8: /cgn2_6/ptodata/1/pubpaa/US08_PUBCOMB.pep:*
9: /cgn2_6/ptodata/1/pubpaa/US09_PUBCOMB.pep:*
10: /cgn2_6/ptodata/1/pubpaa/US09_PUBCOMB.pep:*
11: /cgn2_6/ptodata/1/pubpaa/US09_PUBCOMB.pep:*
12: /cgn2_6/ptodata/1/pubpaa/US09_NEW_PUB.pep:*
13: /cgn2_6/ptodata/1/pubpaa/US10_PUBCOMB.pep:*
14: /cgn2_6/ptodata/1/pubpaa/US10_PUBCOMB.pep:*
15: /cgn2_6/ptodata/1/pubpaa/US10_PUBCOMB.pep:*
16: /cgn2_6/ptodata/1/pubpaa/US10_NEW_PUB.pep:*
17: /cgn2_6/ptodata/1/pubpaa/US60_NEW_PUB.pep:*
18: /cgn2_6/ptodata/1/pubpaa/US60_PUBCOMB.pep:*
```

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	158	100.0	29	US-09-988-851-5	Sequence 5, Appl1
2	148	93.7	27	US-09-988-851-4	Sequence 4, Appl1
3	137	86.7	25	US-09-988-851-3	Sequence 3, Appl1
4	129	81.6	23	US-09-988-851-2	Sequence 2, Appl1
5	92	58.2	20	US-09-988-851-7	Sequence 7, Appl1
6	82	51.9	15	US-09-988-851-6	Sequence 6, Appl1
7	66	41.8	12	US-09-988-851-1	Sequence 1, Appl1
8	55	34.8	12	US-10-108-260A-3540	Sequence 3540, Ap
9	54	34.2	195	US-10-437-963-180535	Sequence 180535,
10	53	33.5	3519	US-09-808-880-4	Sequence 4, Appl1
11	53	33.5	4150	US-09-808-880-2	Sequence 2, Appl1
12	52	32.9	255	US-09-968-067B-10	Sequence 10, Appl
13	52	32.9	255	US-10-335-977-6137	Sequence 6137, Ap
14	52	32.9	264	US-10-335-977-6138	Sequence 6138, Ap
15	51.5	32.6	1081	US-10-437-963-196915	Sequence 196915,

16	51	32.3	81	16	US-10-437-963-118743	Sequence 118743,
17	51	32.3	113	16	US-10-437-963-125727	Sequence 125727,
18	51	32.3	113	16	US-10-767-701-41863	Sequence 41863, A
19	51	32.3	421	15	US-10-369-493-23087	Sequence 23087, A
20	51	32.3	7068	16	US-10-203-293-20	Sequence 20, Appl
21	51	32.3	9477	16	US-10-203-293-37	Sequence 2586, Ap
22	50	31.6	85	12	US-10-276-774-2586	Sequence 151284,
23	50	31.6	113	12	US-10-424-589-151284	Sequence 930, App
24	50	31.6	117	9	US-09-764-866-930	Sequence 19, Appl
25	50	31.6	425	15	US-10-172-585-918	Sequence 22547, A
26	50	31.6	877	15	US-10-369-493-22547	Sequence 3, Appl1
27	50	31.6	3816	15	US-09-808-880-3	Sequence 1101, A
28	49	31.0	280	15	US-10-369-493-11101	Sequence 6140, Ap
29	49	31.0	490	9	US-09-738-626-6140	Sequence 26, Appl
30	49	31.0	1612	16	US-10-647-156-26	Sequence 18, Appl
31	49	31.0	1784	14	US-10-205-032-18	Sequence 62123, A
32	48	30.4	127	12	US-10-425-114-66323	Sequence 66522, A
33	48	30.4	154	12	US-10-425-114-66522	Sequence 2967, Ap
34	48	30.4	165	15	US-10-264-049-2967	Sequence 5998, Ap
35	48	30.4	223	12	US-10-335-977-5998	Sequence 135, App
36	48	30.4	254	10	US-09-768-787-135	Sequence 82, Appl
37	48	30.4	277	10	US-09-988-067B-82	Sequence 5999, Ap
38	48	30.4	277	12	US-10-335-977-5999	Sequence 57083, A
39	48	30.4	378	12	US-10-425-114-57083	Sequence 2, Appl1
40	48	30.4	931	9	US-09-779-081-2	Sequence 75, Appl
41	48	30.4	1609	10	US-09-841-260-75	Sequence 63, Appl
42	48	30.4	1609	13	US-10-007-6893-75	Sequence 75, Appl
43	48	30.4	1609	15	US-10-312-273-85	Sequence 75, Appl
44	48	30.4	1609	16	US-10-762-058-75	Sequence 1035, Ap
45	48	30.4	1617	15	US-10-289-762-1035	

ALIGNMENTS

```
RESULT 1
US-09-988-851-5
; Sequence 5, Application US/0998851
; Publication No. US20020034609A1
; GENERAL INFORMATION:
; APPLICANT: Medical Research Council
; APPLICANT: Marsden, Howard
; APPLICANT: Stow, Nigel
; APPLICANT: McLean, Gordon
; TITLE OF INVENTION: ASSAY FOR SCREENING FOR AN ANTI-VIRAL AGENT
; FILE REFERENCE: P29347A
; CURRENT APPLICATION NUMBER: US/09/998,851
; CURRENT FILING DATE: 2002-03-11
; NUMBER OF SEQ ID NOS: 18
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 5
; LENGTH: 29
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Peptide derived from herpes simplex virus.
US-09-988-851-5

Query Match      100.0%; Score 158; DB 12; Length 29;
Best Local Similarity 100.0%; Pred. No. 1.8e-15;
Matches 29; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      1 VFTGVLAGVWGEGKRVYPPDDKMSFLFA 29
        |||
Db       1 VFTGVLAGVWGEGKRVYPPDDKMSFLFA 29

RESULT 2
US-09-988-851-4
; Sequence 4, Application US/0998851
; Publication No. US20020034609A1
; GENERAL INFORMATION:
; APPLICANT: Medical Research Council
```

APPLICANT: Marsden, Howard
APPLICANT: Stow, Nigel
APPLICANT: McLean, Gordon
TITLE OF INVENTION: ASSAY FOR SCREENING FOR AN ANTI-VIRAL AGENT
FILE REFERENCE: P29347A
CURRENT APPLICATION NUMBER: US/09/998,851
CURRENT FILING DATE: 2002-03-11
NUMBER OF SEQ ID NOS: 18
SOFTWARE: PatentIn version 3.1
SEQ ID NO 4
LENGTH: 27
TYPE: PRT
ORGANISM: Artificial Sequence
FEATURE:
OTHER INFORMATION: Peptide derived from herpes simplex virus.
US-09-998-851-4

Query Match 93.7%; Score 148; DB 12; Length 27;
Best Local Similarity 100.0%; Pred. No. 4.5e-14;
Matches 27; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 3 TGVLAGWGEKGKFPDDXMSFLPA 29
Db 1 TGVLAGWGEKGKFPDDXMSFLPA 27

RESULT 3
US-09-998-851-3
Sequence 3, Application US/09998851
Publication No. US20020034609A1
GENERAL INFORMATION:
APPLICANT: Marsden, Howard
APPLICANT: Stow, Nigel
APPLICANT: McLean, Gordon
TITLE OF INVENTION: ASSAY FOR SCREENING FOR AN ANTI-VIRAL AGENT
FILE REFERENCE: P29347A
CURRENT APPLICATION NUMBER: US/09/998,851
CURRENT FILING DATE: 2002-03-11
NUMBER OF SEQ ID NOS: 18
SOFTWARE: PatentIn version 3.1
SEQ ID NO 3
LENGTH: 25
TYPE: PRT
ORGANISM: Artificial Sequence
FEATURE:
OTHER INFORMATION: Peptide derived from herpes simplex virus.
US-09-998-851-3

Query Match 86.7%; Score 137; DB 12; Length 25;
Best Local Similarity 100.0%; Pred. No. 1.6e-12;
Matches 25; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 5 VLAGWGEKGKFPDDXMSFLPA 29
Db 1 VLAGWGEKGKFPDDXMSFLPA 25

RESULT 4
US-09-998-851-2
Sequence 2, Application US/09998851
Publication No. US20020034609A1
GENERAL INFORMATION:
APPLICANT: Marsden, Howard
APPLICANT: Stow, Nigel
APPLICANT: McLean, Gordon
TITLE OF INVENTION: ASSAY FOR SCREENING FOR AN ANTI-VIRAL AGENT
FILE REFERENCE: P29347A
CURRENT APPLICATION NUMBER: US/09/998,851
CURRENT FILING DATE: 2002-03-11
NUMBER OF SEQ ID NOS: 18
SOFTWARE: PatentIn version 3.1

SEQ ID NO 2
LENGTH: 23
TYPE: PRT
ORGANISM: Artificial Sequence
FEATURE:
OTHER INFORMATION: Peptide derived from herpes simplex virus.
US-09-998-851-2

Query Match 81.6%; Score 129; DB 12; Length 23;
Best Local Similarity 100.0%; Pred. No. 2e-11;
Matches 23; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 7 AGWGEKGKFPDDXMSFLPA 29
Db 1 AGWGEKGKFPDDXMSFLPA 23

RESULT 5
US-09-998-851-7
Sequence 7, Application US/09998851
Publication No. US20020034609A1
GENERAL INFORMATION:
APPLICANT: Marsden, Howard
APPLICANT: Stow, Nigel
APPLICANT: McLean, Gordon
TITLE OF INVENTION: ASSAY FOR SCREENING FOR AN ANTI-VIRAL AGENT
FILE REFERENCE: P29347A
CURRENT APPLICATION NUMBER: US/09/998,851
CURRENT FILING DATE: 2002-03-11
NUMBER OF SEQ ID NOS: 18
SOFTWARE: PatentIn version 3.1
SEQ ID NO 7
LENGTH: 20
TYPE: PRT
ORGANISM: Artificial Sequence
FEATURE:
OTHER INFORMATION: Peptide derived from herpes simplex virus.
US-09-998-851-7

Query Match 58.2%; Score 92; DB 12; Length 20;
Best Local Similarity 100.0%; Pred. No. 3.5e-06;
Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 VFTGVLAGWGEKGKFPV 17
Db 4 VFTGVLAGWGEKGKFPV 20

RESULT 6
US-09-998-851-6
Sequence 6, Application US/09998851
Publication No. US20020034609A1
GENERAL INFORMATION:
APPLICANT: Marsden, Howard
APPLICANT: Stow, Nigel
APPLICANT: McLean, Gordon
TITLE OF INVENTION: ASSAY FOR SCREENING FOR AN ANTI-VIRAL AGENT
FILE REFERENCE: P29347A
CURRENT APPLICATION NUMBER: US/09/998,851
CURRENT FILING DATE: 2002-03-11
NUMBER OF SEQ ID NOS: 18
SOFTWARE: PatentIn version 3.1
SEQ ID NO 6
LENGTH: 15
TYPE: PRT
ORGANISM: Artificial Sequence
FEATURE:
OTHER INFORMATION: Peptide derived from herpes simplex virus.
US-09-998-851-6

Query Match 51.9%; Score 82; DB 12; Length 15;

Best Local Similarity 100.0%; Pred. No. 7e-05;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3 TGVLAGVWGEGKFL 17
Db 1 TGVLAGVWGEGKFL 15

RESULT 7

US-09-998-851-1
; Sequence 1, Application US/09998851
; Publication No. US20020034609a1
; GENERAL INFORMATION:
; APPLICANT: Medical Research Council
; APPLICANT: Marsden, Howard
; APPLICANT: Stow, Nigel
; APPLICANT: McLean, Gordon
; TITLE OF INVENTION: ASSAY FOR SCREENING FOR AN ANTI-VIRAL AGENT
; FILE REFERENCE: P23347A
; CURRENT APPLICATION NUMBER: US/09/998,851
; CURRENT FILING DATE: 2002-03-11
; NUMBER OF SEQ ID NOS: 18
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 1
; LENGTH: 12
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Peptide derived from herpes simplex virus.
US-09-998-851-1

Query Match 41.8%; Score 66; DB 12; Length 12;
Best Local Similarity 100.0%; Pred. No. 0.011;
Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 18 YPFDDKMSFLFA 29
Db 1 YPFDDKMSFLFA 12

RESULT 8

US-10-108-260A-3540
; Sequence 3540, Application US/10108260A
; Publication No. US20040005560a1
; GENERAL INFORMATION:
; APPLICANT: HELIX RESEARCH INSTITUTE
; TITLE OF INVENTION: NO. US20040005560a1e1 full length cDNA
; FILE REFERENCE: H1-A0106
; CURRENT APPLICATION NUMBER: US/10/108,260A
; CURRENT FILING DATE: 2002-03-27
; NUMBER OF SEQ ID NOS: 5458
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 3540
; LENGTH: 182
; TYPE: PRT
; ORGANISM: Homo sapiens
; OTHER INFORMATION: Description of Artificial Sequence: Recombinant
US-10-108-260A-3540

Query Match 34.8%; Score 55; DB 15; Length 182;
Best Local Similarity 50.0%; Pred. No. 7.5;
Matches 10; Conservative 3; Mismatches 7; Indels 0; Gaps 0;

QY 10 WGEQGFYVPPDDKMSFLFA 29
Db 17 WGPQGAAGSAPWDQYTSFLFS 36

RESULT 9

US-10-437-963-180535
; Sequence 180535, Application US/10437963
; Publication No. US20040123343a1
; GENERAL INFORMATION:
; APPLICANT: La Rosa, Thomas J.

APPLICANT: Kovacic, David K.

APPLICANT: Zhou, Yihua

APPLICANT: Cao, Yongwei

APPLICANT: Wu, Wei

APPLICANT: Boukharov, Andrey A.

APPLICANT: Barbazuk, Brad

APPLICANT: Li, Ping

TITLE OF INVENTION: Rice Nucleic Acid Molecules and Other Molecules Associated With

FILE REFERENCE: 38-21(53221)B

CURRENT APPLICATION NUMBER: US/10/437,963

CURRENT FILING DATE: 2003-05-14

NUMBER OF SEQ ID NOS: 204966

SEQ ID NO 180535

LENGTH: 195

TYPE: PRT

ORGANISM: Oryza sativa

FEATURE:

OTHER INFORMATION: Clone ID: PAT_MRT4530_77897C.1.pep

US-10-437-963-180535

Query Match 34.2%; Score 54; DB 16; Length 195;
Best Local Similarity 43.2%; Pred. No. 11;
Matches 16; Conservative 3; Mismatches 6; Indels 12; Gaps 3;

QY 3 TGVLAGVWG--EGKFLV-----YFPD-DKMSFL 27
Db 37 TGTIAGPQGTPEYEGGTVIDRLFGYVPEPMPQPI 73

RESULT 10

US-09-808-880-4
; Sequence 4, Application US/09808880
; Publication No. US20030027287a1
; GENERAL INFORMATION:
; APPLICANT: Betlach, Mary C.
; APPLICANT: Shah, Sanjay Krishnakant
; APPLICANT: McDaniel, Robert
; APPLICANT: Tang, Li
; TITLE OF INVENTION: RECOMBINANT CLEANDOLIDE POLYKETIDE SYNTHASE
; FILE REFERENCE: 30062-20029.00
; CURRENT APPLICATION NUMBER: US/09/808,880
; CURRENT FILING DATE: 2001-03-14
; PRIOR APPLICATION NUMBER: US/09/428,517
; PRIOR FILING DATE: 1999-10-28
; PRIOR APPLICATION NUMBER: 60/120,254
; PRIOR FILING DATE: 1999-02-16
; PRIOR APPLICATION NUMBER: 60/106,100
; PRIOR FILING DATE: 1998-10-29
; NUMBER OF SEQ ID NOS: 12
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 4
; LENGTH: 3519
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: Recombinant
US-09-808-880-4

Query Match 33.5%; Score 53; DB 10; Length 3519;
Best Local Similarity 75.0%; Pred. No. 3.4e+02;
Matches 9; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 7 AGVWGGGKFLV 18
Db 1337 AGVWGGGKFLV 1348

RESULT 11

US-09-808-880-2
; Sequence 2, Application US/09808880
; Publication No. US20030027287a1

GENERAL INFORMATION:
APPLICANT: Betlach, Mary C.
APPLICANT: Shah, Sanjay Krishnakant
APPLICANT: McDaniel, Robert
APPLICANT: Tang, Li
TITLE OF INVENTION: RECOMBINANT CLEANDOLIDE POLYPEPTIDE SYNTHASE
FILE REFERENCE: 30062-20029.00
CURRENT APPLICATION NUMBER: US/09/808,880
CURRENT FILING DATE: 2001-03-14
PRIOR APPLICATION NUMBER: US/09/428,517
PRIOR FILING DATE: 1999-10-28
PRIOR APPLICATION NUMBER: 60/120,254
PRIOR FILING DATE: 1999-02-16
PRIOR APPLICATION NUMBER: 60/106,100
PRIOR FILING DATE: 1998-10-29
NUMBER OF SEQ ID NOS: 12
SOFTWARE: PatentIn Ver. 2.1
SEQ ID NO 2
LENGTH: 4150
TYPE: PRT
ORGANISM: Artificial Sequence
FEATURE:
OTHER INFORMATION: Description of Artificial Sequence: Recombinant
US-09-808-880-2

Query Match 33.5%; Score 53; DB 10; Length 4150;
Best Local Similarity 75.0%; Pred. No. 4.1e+02;
Matches 9; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 7 AGVWGEGGKRVY 18
Db 3838 AGVWGSGGQAVY 3849

RESULT 12
US-09-988-067B-10
Sequence 10, Application US/09988067B
Publication No. US20030124141A1
GENERAL INFORMATION:
APPLICANT: Haas, Rainer
APPLICANT: Kleinhous, Harold
APPLICANT: Tomb, Jean-Francois
APPLICANT: Miller, Charles
APPLICANT: Al-Garawi, Amal
APPLICANT: Odenbreit, Stefan
APPLICANT: Meyer, Thomas
TITLE OF INVENTION: Helicobacter Polypeptides and
TITLE OF INVENTION: Corresponding Polynucleotide Molecules
FILE REFERENCE: 06132/040002
CURRENT APPLICATION NUMBER: US/09/988,067B
CURRENT FILING DATE: 2003-01-31
PRIOR APPLICATION NUMBER: US 08/831,309
PRIOR FILING DATE: 1997-04-01
NUMBER OF SEQ ID NOS: 112
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 10
LENGTH: 255
TYPE: PRT
ORGANISM: Helicobacter pylori
US-09-988-067B-10

Query Match 32.9%; Score 52; DB 10; Length 255;
Best Local Similarity 47.1%; Pred. No. 29;
Matches 8; Conservative 6; Mismatches 3; Indels 0; Gaps 0;

QY 1 VFTGYLAGVWGEGKRV 17
Db 12 LFTGILSGIFGIGGLI 28

RESULT 13
US-10-335-977-6137

Sequence 6137, Application US/10335977
Publication No. US20040052799A1
GENERAL INFORMATION:
APPLICANT: DOUGLAS SMITH et al
TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES
RELATING TO HELICOBLACTER PYLORI FOR
DIAGNOSTICS AND THERAPEUTICS
NUMBER OF SEQUENCES: 10031
CORRESPONDENCE ADDRESS:
ADDRESSEE: LAHIVE & COCKFIELD
STREET: 28 State Street
CITY: Boston
STATE: Massachusetts
COUNTRY: USA
ZIP: 02109-1875
COMPUTER READABLE FORM:
MEDIUM TYPE: CD-ROM ISO9660
COMPUTER: IBM PC Compatible
OPERATING SYSTEM: Windows NT 4.0
SOFTWARE: UNIX
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/10/335,977
FILING DATE: 30-Dec-2002
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/993,002
FILING DATE: 17-Dec-1997
ATTORNEY/AGENT INFORMATION:
NAME: Mandragouras, Amy E.
REGISTRATION NUMBER: 36,207
REFERENCE/DOCKET NUMBER: GTN-018
TELECOMMUNICATION INFORMATION:
TELEPHONE: (617)227-7400
TELEFAX: (617)742-4214
INFORMATION FOR SEQ ID NO: 6137:
SEQUENCE CHARACTERISTICS:
LENGTH: 255 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
HYPOTHETICAL: YES
ORIGINAL SOURCE:
ORGANISM: Helicobacter pylori
FEATURE:
NAME/KEY: misc feature
LOCATION: (R) LOCATION 1...255
SEQUENCE DESCRIPTION: SEQ ID NO: 6137:
US-10-335-977-6137

Query Match 32.9%; Score 52; DB 12; Length 255;
Best Local Similarity 47.1%; Pred. No. 29;
Matches 8; Conservative 6; Mismatches 3; Indels 0; Gaps 0;

QY 1 VFTGYLAGVWGEGKRV 17
Db 12 LFTGILSGIFGIGGLI 28

RESULT 14
US-10-335-977-6138
Sequence 6138, Application US/10335977
Publication No. US20040052799A1
GENERAL INFORMATION:
APPLICANT: DOUGLAS SMITH et al
TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES
RELATING TO HELICOBLACTER PYLORI FOR
DIAGNOSTICS AND THERAPEUTICS
NUMBER OF SEQUENCES: 10031
CORRESPONDENCE ADDRESS:
ADDRESSEE: LAHIVE & COCKFIELD
STREET: 28 State Street
CITY: Boston
STATE: Massachusetts
COUNTRY: USA

ZIP: 02109-1875
 COMPUTER READABLE FORM:
 MEDIUM TYPE: CD-ROM ISO9660
 COMPUTER: IBM PC Compatible
 OPERATING SYSTEM: Windows NT 4.0
 SOFTWARE: UNIX
 CURRENT APPLICATION DATA:
 APPLICATION NUMBER: US/10/335,977
 FILING DATE: 30-Dec-2002
 PRIOR APPLICATION DATA:
 APPLICATION NUMBER: 06/393,002
 FILING DATE: 17-DEC-1997
 ATTORNEY/AGENT INFORMATION:
 NAME: Mandragoras, Amy E.
 REGISTRATION NUMBER: 36,207
 REFERENCE/DOCKET NUMBER: GTN-018
 TELECOMMUNICATION INFORMATION:
 TELEPHONE: (617)227-7400
 TELEFAX: (617)742-4214
 INFORMATION FOR SEQ ID NO: 6138:
 SEQUENCE CHARACTERISTICS:
 LENGTH: 264 amino acids
 TYPE: amino acid
 TOPOLOGY: linear
 MOLECULE TYPE: protein
 HYPOTHETICAL: YES
 ORIGINAL SOURCE:
 ORGANISM: Helicobacter pylori
 FEATURE:
 NAME/KEY: misc feature
 LOCATION: (B) LOCATION 1..264
 SEQUENCE DESCRIPTION: SEQ ID NO: 6138:
 US-10-335-977-6138

Query Match 32.9%; Score 52; DB 12; Length 264;
 Best Local Similarity 47.1%; Pred. No. 30;
 Matches 8; Conservative 6; Mismatches 3; Indels 0; Gaps 0;

QY 1 VFTGVLAGVWGEGKRV 17
 :|||:|:|:|:|:|:|:
 Db 21 LFTGISLGFIRGGMI 37

RESULT 15
 US-10-437-963-196915
 Sequence 196915, Application US/10437963
 Publication No. US20040123343A1
 GENERAL INFORMATION:
 APPLICANT: La Rosa, Thomas J.
 APPLICANT: Kovalic, David K.
 APPLICANT: Zhou, Yihua
 APPLICANT: Cao, Yongwei
 APPLICANT: Wu, Wei
 APPLICANT: Boucharov, Andrey A.
 APPLICANT: Barbazuk, Brad
 APPLICANT: Li, Ping
 TITLE OF INVENTION: Rice Nucleic Acid Molecules and Other Molecules Associated With
 TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement
 FILE REFERENCE: 38-21(53221)B
 CURRENT APPLICATION NUMBER: US/10/437,963
 CURRENT FILING DATE: 2003-05-14
 NUMBER OF SEQ ID NOS: 204966
 SEQ ID NO 196915
 LENGTH: 1081
 TYPE: PRT
 ORGANISM: Oryza sativa
 FEATURE:
 OTHER INFORMATION: Clone ID: PAT_MRT4530_92723C.1.pep
 US-10-437-963-196915

Query Match 32.6%; Score 51.5; DB 16; Length 1081;
 Best Local Similarity 42.9%; Pred. No. 1,66+02;
 Matches 12; Conservative 3; Mismatches 10; Indels 3; Gaps 1;

QY 4 GVLAVGWGEG--GKRVFPFDDKMSPLF 28
 ||| ||| ||| :|:|:|:
 Db 386 GCLAVVWVGWVGEGKRVFSSGNSMKLVF 413
 Search completed: September 30, 2004, 11:33:33
 Job time: 72.9143 secs

This Page Blank (uspto)

Fri Oct 1 12:26:19 2004

us-09-988-851a-5.ra1

Page 1

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: September 30, 2004, 10:58:43 : Search time 20.1619 Seconds
(without alignments)
74.257 Million cell updates/sec

Title: US-09-988-851A-5
Perfect score: 158
Sequence: 1 VFTGVLAGVWGEGKRFVFPDDKMSFLFA 29

Scoring table: BIOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 389414 segs, 51625971 residues

Total number of hits satisfying chosen parameters: 389414

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Issued Patents AA.*
1: /cgn2_6/ptodata/2/iaa/5A.COMB.pep.*
2: /cgn2_6/ptodata/2/iaa/5B.COMB.pep.*
3: /cgn2_6/ptodata/2/iaa/6A.COMB.pep.*
4: /cgn2_6/ptodata/2/iaa/6B.COMB.pep.*
5: /cgn2_6/ptodata/2/iaa/PCTUS.COMB.pep.*
6: /cgn2_6/ptodata/2/iaa/backfilest.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	158	100.0	29	4	US-09-230-405-5
2	148	93.7	27	4	US-09-230-405-4
3	137	86.7	25	4	US-09-230-405-3
4	129	81.6	23	4	US-09-230-405-2
5	92	58.2	20	4	US-09-230-405-7
6	82	51.9	15	4	US-09-230-405-6
7	66	41.8	12	4	US-09-230-405-1
8	60.5	38.3	367	4	US-09-134-000C-6071
9	54.5	34.5	373	4	US-09-107-532A-6660
10	53	33.5	1580	2	US-08-804-227C-11
11	53	33.5	1580	2	US-08-804-198-5
12	53	33.5	1891	2	US-08-804-227C-12
13	53	33.5	1891	2	US-08-804-198-6
14	53	33.5	3519	2	US-09-428-517-4
15	53	33.5	4150	3	US-09-428-517-2
16	51	32.3	4545	2	US-08-804-227C-14
17	51	32.3	4550	2	US-08-804-227C-8
18	51	32.3	4550	2	US-08-804-198-2
19	50	31.6	262	4	US-09-328-352-6552
20	50	31.6	389	3	US-09-307-621-2
21	50	31.6	425	4	US-09-634-955B-19
22	50	31.6	772	1	US-08-258-639A-2
23	50	31.6	772	1	US-08-900-951-2
24	50	31.6	772	5	PCT-US95-07391A-2
25	50	31.6	877	4	US-09-165-396-5
26	50	31.6	877	4	US-09-428-517-3
27	48.5	30.7	539	4	US-09-252-991A-31254

28	48	30.4	371	4	US-09-489-039A-8780	Sequence 8780, Ap
29	48	30.4	931	4	US-09-427-533B-2	Sequence 2, Appl
30	48	30.4	1617	4	US-09-198-452A-1035	Sequence 1035, Ap
31	47	29.7	138	4	US-09-570-921-29	Sequence 29, Appl
32	47	29.7	199	3	US-09-078-317-2	Sequence 2, Appl
33	47	29.7	199	4	US-09-454-818-2	Sequence 2, Appl
34	47	29.7	268	3	US-08-965-056-70	Sequence 70, Appl
35	47	29.7	739	3	US-09-136-251-2	Sequence 2, Appl
36	47	29.7	739	4	US-09-634-496-2	Sequence 2, Appl
37	47	29.7	740	4	US-09-635-145A-2	Sequence 2, Appl
38	46.5	29.4	192	4	US-09-543-681A-5012	Sequence 5012, Ap
39	46.5	29.4	516	3	US-09-197-063-2	Sequence 2, Appl
40	46.5	29.4	903	4	US-09-339-159B-32	Sequence 32, Appl
41	46	29.1	243	4	US-09-540-236-3399	Sequence 3399, Ap
42	46	29.1	322	4	US-09-489-039A-12383	Sequence 12383, A
43	46	29.1	1562	3	US-09-320-878-3	Sequence 3, Appl
44	46	29.1	1562	4	US-09-105-537-35	Sequence 35, Appl
45	46	29.1	1562	4	US-09-141-908-4	Sequence 4, Appl

ALIGNMENTS

RESULT 1
US-09-230-405-5
; Sequence 5, Application US/09230405
; Patent No. 6337074
; GENERAL INFORMATION:
; APPLICANT: Medical Research Council
; TITLE OF INVENTION: ANTI-VIRAL AGENT
; FILE REFERENCE: PI7970C
; CURRENT APPLICATION NUMBER: US/09/230, 405
; CURRENT FILING DATE: 1999-01-25
; NUMBER OF SEQ ID NOS: 18
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 5
; LENGTH: 29
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence:peptide derived
; OTHER INFORMATION: from herpes simplex virus
US-09-230-405-5

Query Match 100.0%; Score 158; DB 4; Length 29;
Best Local Similarity 100.0%; Pred. No. 9e-16;
Matches 29; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 VFTGVLAGVWGEGKRFVFPDDKMSFLFA 29
DB 1 VFTGVLAGVWGEGKRFVFPDDKMSFLFA 29
RESULT 2
US-09-230-405-4
; Sequence 4, Application US/09230405
; Patent No. 6337074
; GENERAL INFORMATION:
; APPLICANT: Medical Research Council
; TITLE OF INVENTION: ANTI-VIRAL AGENT
; FILE REFERENCE: PI7970C
; CURRENT APPLICATION NUMBER: US/09/230, 405
; CURRENT FILING DATE: 1999-01-25
; NUMBER OF SEQ ID NOS: 18
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 4
; LENGTH: 27
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence:peptide derived
; OTHER INFORMATION: from herpes simplex virus
US-09-230-405-4

```
Query Match          93.7%; Score 148; DB 4; Length 27;
Best Local Similarity 100.0%; Pred. No. 2.2e-14;
Matches 27; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3 TGVLAGWGGGKRVYPPDDKMSFLFA 29
DB 1 TGVLAGWGGGKRVYPPDDKMSFLFA 27

RESULT 3
US-09-230-405-3
; Sequence 3, Application US/09230405
; Patent No. 6337074
; GENERAL INFORMATION:
; APPLICANT: Medical Research Council
; TITLE OF INVENTION: ANTI-VIRAL AGENT
; FILE REFERENCE: P17970C
; CURRENT APPLICATION NUMBER: US/09/230,405
; CURRENT FILING DATE: 1999-01-25
; NUMBER OF SEQ ID NOS: 18
; SOFTWARE: Patentin Ver. 2.1
; SEQ ID NO 3
; LENGTH: 25
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence:peptide derived
US-09-230-405-3

Query Match          86.7%; Score 137; DB 4; Length 25;
Best Local Similarity 100.0%; Pred. No. 7.2e-13;
Matches 25; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 5 VLAGWGGGKRVYPPDDKMSFLFA 29
DB 1 VLAGWGGGKRVYPPDDKMSFLFA 25

RESULT 4
US-09-230-405-2
; Sequence 2, Application US/09230405
; Patent No. 6337074
; GENERAL INFORMATION:
; APPLICANT: Medical Research Council
; TITLE OF INVENTION: ANTI-VIRAL AGENT
; FILE REFERENCE: P17970C
; CURRENT APPLICATION NUMBER: US/09/230,405
; CURRENT FILING DATE: 1999-01-25
; NUMBER OF SEQ ID NOS: 18
; SOFTWARE: Patentin Ver. 2.1
; SEQ ID NO 2
; LENGTH: 23
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence:peptide derived
US-09-230-405-2

Query Match          81.6%; Score 129; DB 4; Length 23;
Best Local Similarity 100.0%; Pred. No. 9e-12;
Matches 23; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 7 AGWGGGKRVYPPDDKMSFLFA 29
DB 1 AGWGGGKRVYPPDDKMSFLFA 23

RESULT 5
US-09-230-405-7
; Sequence 7, Application US/09230405
```

```
; Patent No. 6337074
; GENERAL INFORMATION:
; APPLICANT: Medical Research Council
; TITLE OF INVENTION: ANTI-VIRAL AGENT
; FILE REFERENCE: P17970C
; CURRENT APPLICATION NUMBER: US/09/230,405
; CURRENT FILING DATE: 1999-01-25
; NUMBER OF SEQ ID NOS: 18
; SOFTWARE: Patentin Ver. 2.1
; SEQ ID NO 7
; LENGTH: 20
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence:peptide derived
US-09-230-405-7

Query Match          58.2%; Score 92; DB 4; Length 20;
Best Local Similarity 100.0%; Pred. No. 1.3e-06;
Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 VFTGVLAGWGGGKRV 17
DB 4 VFTGVLAGWGGGKRV 20

RESULT 6
US-09-230-405-6
; Sequence 6, Application US/09230405
; Patent No. 6337074
; GENERAL INFORMATION:
; APPLICANT: Medical Research Council
; TITLE OF INVENTION: ANTI-VIRAL AGENT
; FILE REFERENCE: P17970C
; CURRENT APPLICATION NUMBER: US/09/230,405
; CURRENT FILING DATE: 1999-01-25
; NUMBER OF SEQ ID NOS: 18
; SOFTWARE: Patentin Ver. 2.1
; SEQ ID NO 6
; LENGTH: 15
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence:peptide derived
US-09-230-405-6

Query Match          51.9%; Score 82; DB 4; Length 15;
Best Local Similarity 100.0%; Pred. No. 2.6e-05;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3 TGVLAGWGGGKRV 17
DB 1 TGVLAGWGGGKRV 15

RESULT 7
US-09-230-405-1
; Sequence 1, Application US/09230405
; Patent No. 6337074
; GENERAL INFORMATION:
; APPLICANT: Medical Research Council
; TITLE OF INVENTION: ANTI-VIRAL AGENT
; FILE REFERENCE: P17970C
; CURRENT APPLICATION NUMBER: US/09/230,405
; CURRENT FILING DATE: 1999-01-25
; NUMBER OF SEQ ID NOS: 18
; SOFTWARE: Patentin Ver. 2.1
; SEQ ID NO 1
; LENGTH: 12
; TYPE: PRT
; ORGANISM: Artificial Sequence
```


FEATURE:
OTHER INFORMATION: Description of Artificial Sequence:peptide derived
OTHER INFORMATION: from herpes simplex virus
US-09-230-405-1

Query Match 41.8%; Score 66; DB 4; Length 12;
Best Local Similarity 100.0%; Pred. No. 0.0037;
Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 18 YPFDDKMSFLFA 29
Db 1 YPFDDKMSFLFA 12

RESULT 8
US-09-134-000C-6071
Sequence 6071, Application US/09134000C
Patent No. 6617156
GENERAL INFORMATION:
APPLICANT: Lynn Doucette-Stamm et al
TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO
FILE REFERENCE: 032796-0032
CURRENT APPLICATION NUMBER: US/09/134.000C
CURRENT FILING DATE: 1998-08-13
PRIOR APPLICATION NUMBER: US 60/055,778
PRIOR FILING DATE: 1997-08-15
NUMBER OF SEQ ID NOS: 6812
SOFTWARE: PatentIn version 3.1
SEQ ID NO 6071
LENGTH: 367
TYPE: PRT
ORGANISM: Enterococcus faecalis
US-09-134-000C-6071

Query Match 38.3%; Score 60.5; DB 4; Length 367;
Best Local Similarity 57.1%; Pred. No. 0.82;
Matches 12; Conservative 2; Mismatches 6; Indels 1; Gaps 1;

QY 9 VMGEGKRVYPPD-DKMSFLF 28
Db 327 VMGEGKRVYPPDKMDLEFNF 347

RESULT 9
US-09-107-532A-6660
Sequence 6660, Application US/09107532A
Patent No. 6583275

GENERAL INFORMATION:
APPLICANT: Lynn A Doucette-Stamm and David Bush
TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO
ENTEROCOCCUS FAECIUM FOR DIAGNOSTICS AND THERAPEUTICS
NUMBER OF SEQUENCES: 7310
CORRESPONDENCE ADDRESS:
ADDRESSEE: GENOME THERAPEUTICS CORPORATION
STREET: 100 Beaver Street
CITY: Waltham
STATE: Massachusetts
COUNTRY: USA
ZIP: 02354
COMPUTER READABLE FORM:
MEDIUM TYPE: CD-ROM ISO9660
COMPUTER: PC
OPERATING SYSTEM: <Unknown>
SOFTWARE: ASCII
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/107,532A
FILING DATE: 30-Jun-1998
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 60/085,598
FILING DATE: 14 May 1998
APPLICATION NUMBER: 60/051571
FILING DATE: July 2, 1997

ATTORNEY/AGENT INFORMATION:
NAME: Ariniello, Pamela Deneke
REGISTRATION NUMBER: 40,489
REFERENCE/DOCKET NUMBER: GTC-012
TELECOMMUNICATION INFORMATION:
TELEPHONE: (781)893-5007
TELEFAX: (781)893-8277

INFORMATION FOR SEQ ID NO: 6660:
SEQUENCE CHARACTERISTICS:
LENGTH: 373 amino acids
TYPE: amino acid

MOLECULE TYPE: linear
HYPOTHETICAL: YES
ORIGINAL SOURCE:
ORGANISM: Enterococcus faecium

FEATURE:
NAME/KEY: misc feature
LOCATION: (3) LOCATION 1..373
SEQUENCE DESCRIPTION: SEQ ID NO: 6660:
US-09-107-532A-6660

Query Match 34.5%; Score 54.5; DB 4; Length 373;
Best Local Similarity 47.6%; Pred. No. 5.9;
Matches 10; Conservative 4; Mismatches 6; Indels 1; Gaps 1;

QY 9 VMGEGKRVYPPD-DKMSFLF 28
Db 333 VMGEGKRVYPTNKDLEFNF 353

RESULT 10
US-08-804-227C-11
Sequence 11, Application US/08804227C
Patent No. 5876991
GENERAL INFORMATION:
APPLICANT: Dehoff, Bradley S.
APPLICANT: Kubstoss, Stuart A.
APPLICANT: Rostock, Paul R., Jr.
TITLE OF INVENTION: POLYPEPTIDE SYNTHASE GENES
NUMBER OF SEQUENCES: 15
CORRESPONDENCE ADDRESS:
ADDRESSEE: THOMAS G. PLANT 1501
STREET: LILLY CORPORATE CENTER
CITY: INDIANAPOLIS
STATE: IN
COUNTRY: USA
ZIP: 46285
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM compatible
OPERATING SYSTEM: MS-DOS
SOFTWARE: ASCII(DOS) text only
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/804,227C
FILING DATE: February 21, 1997
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: Plant, Thomas, G.
REGISTRATION NUMBER: 35,784
REFERENCE/DOCKET NUMBER: X-8231
TELECOMMUNICATION INFORMATION:
TELEPHONE: 317-276-2459
INFORMATION FOR SEQ ID NO: 11:
SEQUENCE CHARACTERISTICS:
LENGTH: 1580 amino acids
TYPE: amino acid
MOLECULE TYPE: peptide
US-08-804-227C-11

Query Match 33.5%; Score 53; DB 2; Length 1580;

Best Local Similarity 75.0%; Pred. No. 44;
Matches 9; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 7 AGWGGGKFPY 18
Db 1287 AGWGGGQSVY 1298

RESULT 11
US-08-804-198-5

; Sequence 5, Application US/08804198
; Patent No. 5945320

GENERAL INFORMATION:

APPLICANT: Burgett, Stanley G.

APPLICANT: Khuskose, Stuart A.

APPLICANT: Rao, Nagaraia R.

APPLICANT: Richardson, Mark A.

APPLICANT: Rosteck, Paul R., Jr.

TITLE OF INVENTION: PLATENOIDE SYNTHASE GENE

NUMBER OF SEQUENCES: 6

CORRESPONDENCE ADDRESS:

ADDRESSEE: PAUL R. CANTRELL 1138

STREET: LILLY CORPORATE CENTER

CITY: INDIANAPOLIS

STATE: IN

COUNTRY: USA

ZIP: 46285

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

COMPUTER: Macintosh

OPERATING SYSTEM: Macintosh 7.0

SOFTWARE: Microsoft Word 5.1

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/804,198

FILING DATE:

CLASSIFICATION: 435

ATTORNEY/AGENT INFORMATION:

NAME: CANTRELL, PAUL R.

REGISTRATION NUMBER: 36,470

REFERENCE/DOCKET NUMBER: P9113

TELECOMMUNICATION INFORMATION:

TELEPHONE: 317-276-3885

INFORMATION FOR SEQ ID NO: 5:

SEQUENCE CHARACTERISTICS:

LENGTH: 1580 amino acids

TYPE: amino acid

TOPOLOGY: unknown

MOLECULE TYPE: peptide

US-08-804-198-5

Query Match 33.5%; Score 53; DB 2; Length 1580;

Best Local Similarity 75.0%; Pred. No. 44;
Matches 9; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 7 AGWGGGKFPY 18
Db 1287 AGWGGGQSVY 1298

RESULT 12

US-08-804-227C-12

; Sequence 12, Application US/08804227C

; Patent No. 5878391

GENERAL INFORMATION:

APPLICANT: Dehoff, Bradley S.

APPLICANT: Khuskose, Stuart A.

APPLICANT: Rosteck, Paul R., Jr.

APPLICANT: Sutton, Kimberly L.

TITLE OF INVENTION: POLYKETIDE SYNTHASE GENES

NUMBER OF SEQUENCES: 15

CORRESPONDENCE ADDRESS:

ADDRESSEE: THOMAS G. PLANT 1501

STREET: LILLY CORPORATE CENTER

CITY: INDIANAPOLIS

STATE: IN

COUNTRY: USA

ZIP: 46285

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

COMPUTER: IBM Compatible

OPERATING SYSTEM: MS-DOS

SOFTWARE: ASCII (DOS) Text only

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/804,227C

FILING DATE: February 21, 1997

CLASSIFICATION: 435

ATTORNEY/AGENT INFORMATION:

NAME: Plant, Thomas, G.

REGISTRATION NUMBER: 35,784

REFERENCE/DOCKET NUMBER: X-8231

TELECOMMUNICATION INFORMATION:

TELEPHONE: 317-276-2459

INFORMATION FOR SEQ ID NO: 12:

SEQUENCE CHARACTERISTICS:

LENGTH: 1891 amino acids

TYPE: amino acid

TOPOLOGY: unknown

MOLECULE TYPE: peptide

US-08-804-227C-12

Query Match 33.5%; Score 53; DB 2; Length 1891;

Best Local Similarity 75.0%; Pred. No. 53;

Matches 9; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 7 AGWGGGKFPY 18
Db 1372 AGWGGGQSVY 1383

RESULT 13

US-08-804-198-6

; Sequence 6, Application US/08804198

; Patent No. 5945320

GENERAL INFORMATION:

APPLICANT: Burgett, Stanley G.

APPLICANT: Khuskose, Stuart A.

APPLICANT: Rao, Nagaraia R.

APPLICANT: Richardson, Mark A.

APPLICANT: Rosteck, Paul R., Jr.

TITLE OF INVENTION: PLATENOIDE SYNTHASE GENE

NUMBER OF SEQUENCES: 6

CORRESPONDENCE ADDRESS:

ADDRESSEE: PAUL R. CANTRELL 1138

STREET: LILLY CORPORATE CENTER

CITY: INDIANAPOLIS

STATE: IN

COUNTRY: USA

ZIP: 46285

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

COMPUTER: Macintosh

OPERATING SYSTEM: Macintosh 7.0

SOFTWARE: Microsoft Word 5.1

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/804,198

FILING DATE:

CLASSIFICATION: 435

ATTORNEY/AGENT INFORMATION:

NAME: CANTRELL, PAUL R.

REGISTRATION NUMBER: 36,470

REFERENCE/DOCKET NUMBER: P9113

TELECOMMUNICATION INFORMATION:

TELEPHONE: 317-276-3885

INFORMATION FOR SEQ ID NO: 6:

SEQUENCE CHARACTERISTICS:

LENGTH: 1891 amino acids

Fri Oct 1 12:26:19 2004

us-09-988-851a-5.ral

Page 5

TYPE: amino acid
TOPOLOGY: unknown
MOLECULE TYPE: peptide
US-08-804-198-6

Query Match 33.5%; Score 53; DB 2; Length 1891;
Best Local Similarity 75.0%; Pred. No. 53;
Matches 9; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

OY 7 AGVWEGGKFEY 18
Db 1372 AGVWGGGQAVY 1383

RESULT 14
US-09-428-517-4
Sequence 4, Application US/09428517
Parent No. 6251636
GENERAL INFORMATION:
APPLICANT: Betlach, Mary C.
APPLICANT: Shah, Sanjay Krishnakant
APPLICANT: McDaniel, Robert
APPLICANT: Tang, Li
TITLE OF INVENTION: RECOMBINANT OLEANDOLIDE POLYKETIDE SYNTHASE
FILE REFERENCE: 30062-20029.00
CURRENT APPLICATION NUMBER: US/09/428,517
CURRENT FILING DATE: 1999-10-28
EARLIER APPLICATION NUMBER: 60/120,254
EARLIER FILING DATE: 1999-02-16
EARLIER APPLICATION NUMBER: 60/106,100
EARLIER FILING DATE: 1998-10-29
NUMBER OF SEQ ID NOS: 12
SOFTWARE: PatentIn Ver. 2.1
SEQ ID NO 4
LENGTH: 3519
TYPE: PRT
ORGANISM: Artificial Sequence
FEATURE:
OTHER INFORMATION: Description of Artificial Sequence: Recombinant
OTHER INFORMATION: Oleandolide PKS
US-09-428-517-4

Query Match 33.5%; Score 53; DB 3; Length 3519;
Best Local Similarity 75.0%; Pred. No. 1e+02;
Matches 9; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

OY 7 AGVWEGGKFEY 18
Db 1337 AGVWGGGQAVY 1348

RESULT 15
US-09-428-517-2
Sequence 2, Application US/09428517
Parent No. 6251636
GENERAL INFORMATION:
APPLICANT: Betlach, Mary C.
APPLICANT: Shah, Sanjay Krishnakant
APPLICANT: McDaniel, Robert
APPLICANT: Tang, Li
TITLE OF INVENTION: RECOMBINANT OLEANDOLIDE POLYKETIDE SYNTHASE
FILE REFERENCE: 30062-20029.00
CURRENT APPLICATION NUMBER: US/09/428,517
CURRENT FILING DATE: 1999-10-28
EARLIER APPLICATION NUMBER: 60/120,254
EARLIER FILING DATE: 1999-02-16
EARLIER APPLICATION NUMBER: 60/106,100
EARLIER FILING DATE: 1998-10-29
NUMBER OF SEQ ID NOS: 12
SOFTWARE: PatentIn Ver. 2.1
SEQ ID NO 2
LENGTH: 4150
TYPE: PRT

ORGANISM: Artificial Sequence
FEATURE:
OTHER INFORMATION: Description of Artificial Sequence: Recombinant
OTHER INFORMATION: Oleandolide PKS
US-09-428-517-2

Query Match 33.5%; Score 53; DB 3; Length 4150;
Best Local Similarity 75.0%; Pred. No. 1.2e+02;
Matches 9; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

OY 7 AGVWEGGKFEY 18
Db 3838 AGVWGGGQAVY 3849

Search completed: September 30, 2004, 11:00:05
Job time: 21.1619 secs

Page Blank (uspto)

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: September 30, 2004, 10:58:43 ; Search time 61.0381 Seconds

(without alignments)
149,907 Million cell updates/sec

Title: US-09-988-851A-5

Perfect score: 158
Sequence: 1 VFTGVLAGVWEGGKRVPPDDXKMSFLFA 29

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 1017041 seqs, 315518202 residues

Total number of hits satisfying chosen parameters: 1017041

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%

Listing first 45 summaries

Database : SPTREMBL_25:*

1: sp_archaea:*

2: sp_bacteria:*

3: sp_fungi:*

4: sp_human:*

5: sp_invertebrate:*

6: sp_mammal:*

7: sp_mhc:*

8: sp_organelle:*

9: sp_phage:*

10: sp_plant:*

11: sp_rodent:*

12: sp_virus:*

13: sp_vertebrate:*

14: sp_unclassified:*

15: sp_virus:*

16: sp_bacteriaph:*

17: sp_archaeap:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	92	58.2	761	12 Q806C1	Q806C1 simian herp
2	60.5	38.3	355	16 Q83615	Q83615 enterococcu
3	57	36.1	637	3 Q9P8P2	Q9P8P2 cryptococcu
4	56	35.4	226	16 Q67094	Q67094 aquifex aeo
5	55	34.8	192	4 Q8N8V3	Q8N8V3 homo sapien
6	55	34.8	634	3 Q8X024	Q8X024 cryptococcu
7	55	34.8	634	3 Q96W00	Q96W00 cryptococcu
8	54	34.2	195	10 Q8S1Y5	Q8S1Y5 oryza sativ
9	54	34.2	310	16 Q8D7N6	Q8D7N6 streptococc
10	54	34.2	1087	5 Q9BMG3	Q9BMG3 leishmania
11	53	33.5	294	16 Q7UEN7	Q7UEN7 rhodospirell
12	53	33.5	854	15 Q97016	Q97016 human immun
13	53	33.5	1151	11 Q9J130	Q9J130 rattus norv
14	53	33.5	4150	2 Q9K1V4	Q9K1V4 streptomyce
15	52.5	33.2	355	16 Q8RD34	Q8RD34 thermoaer
16	52	32.9	255	16 Q25388	Q25388 helicobacte

17	52	32.9	255	16 Q9ZLGI	Q9ZLGI helicobacte
18	52	32.9	270	17 Q8PT69	Q8PT69 methanosarc
19	52	32.9	709	10 Q8LPL9	Q8LPL9 arabidopsis
20	52	32.9	721	10 Q9ZUY2	Q9ZUY2 arabidopsis
21	51.5	32.6	1485	2 Q84IT5	Q84IT5 streptomyce
22	51	32.3	113	10 Q8S1I4	Q8S1I4 oryza sativ
23	51	32.3	281	16 Q87CE6	Q87CE6 xyloella fas
24	51	32.3	363	16 Q9CWC2	Q9CWC2 pasteurella
25	51	32.3	363	16 Q7VKM5	Q7VKM5 haemophilus
26	51	32.3	848	16 Q8A3P6	Q8A3P6 bacteroides
27	51	32.3	9477	2 Q914X3	Q914X3 streptomyce
28	50.5	32.0	442	16 Q93R29	Q93R29 streptomyce
29	50.5	32.0	449	12 Q91R77	Q91R77 chilo iride
30	50	31.6	205	4 Q9H628	Q9H628 homo sapien
31	50	31.6	276	16 Q98NE3	Q98NE3 rhizobium 1
32	50	31.6	279	2 Q934L5	Q934L5 vibrio salm
33	50	31.6	389	2 Q9FDE0	Q9FDE0 bacillus st
34	50	31.6	574	17 Q8TKX6	Q8TKX6 methanosarc
35	50	31.6	668	16 Q8A1I7	Q8A1I7 bacteroides
36	50	31.6	877	3 Q06593	Q06593 saccharomyc
37	50	31.6	322	2 Q94873	Q94873 lyobacter
38	50	31.6	3816	2 Q9K1V3	Q9K1V3 streptomyce
39	49	31.0	118	16 Q92LE0	Q92LE0 rhizobium m
40	49	31.0	156	16 Q9KCG3	Q9KCG3 bacillus ha
41	49	31.0	240	5 Q18527	Q18527 caenorhabdi
42	49	31.0	251	16 Q9XBB8	Q9XBB8 straphylococ
43	49	31.0	251	16 Q8NY77	Q8NY77 straphylococ
44	49	31.0	350	16 Q8YQ02	Q8YQ02 bacteroides
45	49	31.0	490	16 Q8NN09	Q8NN09 corynebacte

ALIGNMENTS

RESULT 1	ID	Q806C1	PRELIMINARY;	PRT;	761 AA.
AC	Q806C1				
DT	01-JUN-2003	(TREMBLrel. 24, Created)			
DT	01-JUN-2003	(TREMBLrel. 25, Last sequence update)			
DT	01-OCT-2003	(TREMBLrel. 25, Last annotation update)			
DE	Component of DNA helicase-primase complex (Component of				
DE	helicase-primase complex).				
GN	Uls.				
OS	Simian herpes B virus (Cercopithecoid herpesvirus 1) (sbv).				
OC	Viruses: dsDNA viruses, no RNA stage; Herpesviridae;				
OC	Alphaherpesvirinae; Simplexvirus.				
OX	NCBI_TaxID=10325;				
RN	(1)				
RP	SEQUENCE FROM N.A.				
RC	STRATIN=E2490;				
RX	MEDLINE=22607624; PubMed=12743273;				
RA	Ohsawa K., Black D.H., Sato H., Rogers K., Eberle R.;				
RT	"Complete Sequence and Comparative Analysis of Herpes B				
RT	virus (Cercopithecine Herpesvirus 1) genome and comparison with the UL				
RT	region of other primate herpesviruses.";				
RL	Arch. Virol. 148:989-997(2003).				
RN	(2)				
RP	SEQUENCE FROM N.A.				
RC	STRATIN=E2490;				
RX	MEDLINE=22628476; PubMed=12743273;				
RA	Pereleygina L., Zhu L., Zurkhlen H., Mills R., Borodovsky M.;				
RT	Hilliard J.K.;				
RT	"Complete Sequence and Comparative Analysis of Herpes B				
RT	virus (Cercopithecine Herpesvirus 1) from a Rhesus Monkey.";				
RL	J. Virol. 77:6167-6177(2003).				
DR	EMBL; AB096160; BAC58047.1; -				
DR	EMBL; AF533768; AAP41426.1; -				
DR	GO; GO:0004386; F:helicase activity; IEA.				
DR	GO; GO:0019075; P:Viral genome replication; IEA.				
DR	InterPro; IPR004996; Herpes_HEPA.				
DR	Pfam; PF03324; Herpes_HEPA; 1.				
KM	Helicase.				
SC	SEQUENCE 761 AA; 79728 MW; 9AD9F9B6C6CF8915E CRC64;				

Query Match 58.2%; Score 92; DB 12; Length 761;
 Best Local Similarity 53.6%; Pred. No. 0.00018;
 Matches 15; Conservative 5; Mismatches 8; Indels 0; Gaps 0;

OY 1 VFTGVLGVWGEKGKFPYPPDDKMSFLF 26
 DB 723 VFRGALTGLWGECCGCFAYPPERRSLFWF 750

RESULT 2

ID O83616 PRELIMINARY; PRT; 355 AA.

AC O83616: 01-JUN-2003 (TREMBlrel. 24, Created)
 DT 01-JUN-2003 (TREMBlrel. 24, Last sequence update)
 DT 01-JUN-2003 (TREMBlrel. 24, Last annotation update)
 DE Hypothetical protein.
 GN EF1125.
 OS Enterococcus faecalis (Streptococcus faecalis).
 OC Bacteria; Firmicutes; Lactobacillales; Enterococcaceae; Enterococcus.
 OX NCBI_TaxID=1351;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=V583 / ATCC 700802;
 RX MEDLINE=22550857; PubMed=12663927;
 RA Paulsen I.T., Banerjee L., Myers G.S.A., Nelson K.B., Seshadri R.,
 RA Read T.D., Fouts D.E., Eisen J.A., Gill S.R., Heidelberg J.F.,
 RA Detlekin H., Dodson R.J., Umayam L., Brinkac L., Beanan M.,
 RA Daugherty S., DeBoy R.T., Durkin S., Kolonay J., Madupu R., Nelson W.,
 RA Vamathevan J., Tran B., Upton J., Hansen T., Shetty J., Khouri H.,
 RA Ureback T., Radune D., Ketchum K.A., Dougherty B.A., Fraser C.M.;
 RT "Role of mobile DNA in the evolution of vancomycin-resistant
 RT Enterococcus faecalis";
 RL Science 289:2071-2074(2003).
 DR EMBL; AE016950; AAC0925.1; -.
 DR TIGR; EF1125; -.
 KW Hypothetical protein; Complete proteome.
 SQ SEQUENCE 355 AA; 40590 MW; 8787B4BBA466B4 CRC64;

Query Match 38.3%; Score 60.5; DB 16; Length 355;
 Best Local Similarity 57.1%; Pred. No. 2.5;
 Matches 12; Conservative 2; Mismatches 6; Indels 1; Gaps 1;

OY 9 VMCGEKGKFPYPPDDKMSFLF 28
 DB 315 IMQVGKFPYPPDDKXNLEPNF 335

RESULT 3

ID O9P8P2 PRELIMINARY; PRT; 637 AA.

AC O9P8P2: 01-OCT-2000 (TREMBlrel. 15, Created)
 DT 01-OCT-2000 (TREMBlrel. 15, Last sequence update)
 DT 01-JUN-2003 (TREMBlrel. 24, Last annotation update)
 DE Secreted phospholipase B.
 GN PLB1.
 OS Cryptococcus neoformans var. neoformans.
 OC Eukaryota; Fungi; Basidiomycota; Hymenycetes; Heterobasidiomycetes;
 OC Tremellomycetidae; Tremellales; Tremellaceae; Filobasidiella.
 OX NCBI_TaxID=40410;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=H99;
 RX MEDLINE=20572097; PubMed=11123698;
 RA Cox G.M., McPade H.C., Chen S.C.A., Tucker S.C., Gottfredsson M.,
 RA Wright L.C., Sorrell T.C., Leidich S.D., Casadevall A., Channoum M.A.,
 RA Perfect J.R.;
 RT "Extracellular phospholipase activity is a virulence factor for
 RT Cryptococcus neoformans";
 RL Mol. Microbiol. 39:166-175(2001).
 DR EMBL; AF223383; AAF65220.1; -.

DR GO; GO:0004620; F:phospholipase activity; IEA.
 DR GO; GO:0003395; F:phospholipid catabolism; IEA.
 DR InterPro; IPR002642; PLAC.
 DR Pfam; PF01735; PLAC_B; 1.
 DR SMART; SM00022; PLAC; 1.
 SQ SEQUENCE 637 AA; 68789 MW; C37B979D69A330 CRC64;

Query Match 36.1%; Score 57; DB 3; Length 637;
 Best Local Similarity 41.7%; Pred. No. 15;
 Matches 10; Conservative 4; Mismatches 10; Indels 0; Gaps 0;

OY 3 TGVLAGVWGEKGKFPYPPDDKMSF 26
 DB 164 TTVLENTMWNIDSNLVFPDDKLSF 187

RESULT 4

ID O67094 PRELIMINARY; PRT; 226 AA.

AC O67094: 01-AUG-1998 (TREMBlrel. 07, Created)
 DT 01-AUG-1998 (TREMBlrel. 07, Last sequence update)
 DT 01-JUN-2003 (TREMBlrel. 24, Last annotation update)
 DE Hypothetical protein AQ_963.
 GN AQ_963.
 OS Aquifex aeolicus.
 OC Bacteria; Aquificae; Aquificales; Aquificaceae; Aquifex.
 OX NCBI_TaxID=63363;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=VF5;
 RX MEDLINE=98196666; PubMed=9537320;
 RA Decker G., Warren P.V., Gaasterland T., Young W.G., Lenox A.L.,
 RA Graham D.E., Overbeek R., Sneed M.A., Keller M., Aubay M., Huber R.,
 RA Feldman R.A., Short J.M., Olson G.J., Shanon R.V.;
 RT "The complete genome of the hyperthermophilic bacterium Aquifex
 RT aeolicus";
 RL Nature 392:353-358(1998).
 DR EMBL; AE000715; AAC07057.1; -.
 DR PIR; C70383; C70383.
 KW Hypothetical protein; Complete proteome.
 SQ SEQUENCE 226 AA; 25939 MW; 29D45D227634281B CRC64;

Query Match 35.4%; Score 56; DB 16; Length 226;
 Best Local Similarity 50.0%; Pred. No. 6.7;
 Matches 9; Conservative 4; Mismatches 5; Indels 0; Gaps 0;

OY 1 VFTGVLGVWGEKGKFPY 18
 DB 164 IFTGLANNAGGNTYW 181

RESULT 5

ID O8N8V3 PRELIMINARY; PRT; 182 AA.

AC O8N8V3: 01-OCT-2002 (TREMBlrel. 22, Created)
 DT 01-OCT-2002 (TREMBlrel. 22, Last sequence update)
 DT 01-OCT-2002 (TREMBlrel. 22, Last annotation update)
 DE Hypothetical protein FLU38820.
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Homidae; Homo.
 OX NCBI_TaxID=9606;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC TISSUE=Liver;
 RA Oshima A., Takahashi-Fujii A., Tanase T., Imose N., Takeuchi K.,
 RA Arita M., Musashino K., Yuki H., Hara H., Sugiyama T., Irie R.,
 RA Otsuki T., Sato H., Ota T., Makamatsu A., Ishii S., Yamamoto J.,
 RA Isono Y., Kawai-Hio Y., Saito K., Nishikawa T., Kimura K.,
 RA Yamashita H., Matsuo K., Nakamura Y., Sekine M., Kikuchi H., Kanda K.,
 RA Wagaatsuma M., Murakawa K., Kanehori K., Sugiyama A., Kawakami B.,

RA Suzuki Y., Sugano S., Nagahari K., Masuno Y., Nagai K., Isogai T.;
 RT "EDO human cDNA sequencing project."
 RL Submitted (JUL-2002) to the EMBL/GenBank/DBJ databases.
 DR EMBL; AK096139; BAC04708.1; -.
 KW Hypothetical protein.
 SQ SEQUENCE 182 AA; 20264 MW; 1F7F88460A324A9C CRC64;
 QY
 Db 10 WGEKGKVFYPPDDKMSFLPA 29
 17 WPGGAGSAPWDQTVSFLFS 36

RESULT 6
 Q8X0Z4 PRELIMINARY; PRT; 634 AA.
 AC Q8X0Z4; (TREMREL. 20, Created)
 DT 01-MAR-2002 (TREMREL. 20, Last sequence update)
 DT 01-JUN-2003 (TREMREL. 24, Last annotation update)
 DE Phospholipase B.
 GN PLB1.
 OS Cryptococcus bacillisporus (Filobasidiella neoformans var. bacillisporus).
 OC Eukaryota; Fungi; Basidiomycota; Hymenomycetes; Heterobasidiomycetes;
 OC Tremellomycetidae; Tremellales; Tremellaceae; Filobasidiella.
 OX NCBI_TaxID=37769;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=MM27055; Sorrell T.C., Meyer W.;
 AC Latouche G.N., Sorrell T.C., Meyer W.;
 RT "Isolation and Characterisation of the phospholipase B gene of
 RT Cryptococcus neoformans var gattii Serotype C."
 RL Submitted (DEC-2000) to the EMBL/GenBank/DBJ databases.
 DR EMBL; AJ302038; CAC83081.1; -.
 DR GO; GO:0004620; F:phospholipase activity; IEA.
 DR GO; GO:0009395; P:phospholipid catabolism; IEA.
 DR InterPro; IPR002642; PLAC.
 DR Pfam; PF01735; PLA2_B; 1.
 DR SMART; SM00022; PLAC; 1.
 SQ SEQUENCE 634 AA; 68187 MW; 19249B2250190239 CRC64;
 QY
 Db 3 TGVLAGWGEKGKVFYPPDDKMSF 26
 164 TSLLENLWNIDSNLIFPDDKVSF 187

Query March 34.8%; Score 55; DB 3; Length 634;
 Best Local Similarity 37.5%; Pred. No. 29;
 Matches 9; Conservative 5; Mismatches 10; Indels 0; Gaps 0;

RESULT 7
 Q96MX0 PRELIMINARY; PRT; 634 AA.
 AC Q96MX0;
 DT 01-DEC-2001 (TREMREL. 19, Created)
 DT 01-DEC-2001 (TREMREL. 19, Last sequence update)
 DT 01-JUN-2003 (TREMREL. 24, Last annotation update)
 DE Phospholipase B (EC 3.1.1.5).
 GN PLB OR PLB1.
 OS Cryptococcus bacillisporus (Filobasidiella neoformans var. bacillisporus).
 OC Eukaryota; Fungi; Basidiomycota; Hymenomycetes; Heterobasidiomycetes;
 OC Tremellomycetidae; Tremellales; Tremellaceae; Filobasidiella.
 OX NCBI_TaxID=37769;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=MM276;
 RA Latouche G.N., Sorrell T.C., Cox G.M., Perfect J.P., Meyer W.;
 RT "Isolation and characterisation of the phospholipase b gene of

RT Cryptococcus neoformans var. gattii."
 RL Submitted (MAY-1999) to the EMBL/GenBank/DBJ databases.
 RN [2]
 RP SEQUENCE FROM N.A.
 RC STRAIN=mm 276;
 RA Latouche G.N., Sorrell T.C., Meyer W.;
 RT "Isolation and sequencing of the cDNA transcript of the phospholipase
 RT B gene from Cryptococcus neoformans var. gattii (serotype B)."
 RL Submitted (DEC-2000) to the EMBL/GenBank/DBJ databases.
 DR EMBL; AJ238508; CAB90748.2; -.
 DR EMBL; AJ302039; CAC83082.1; -.
 DR GO; GO:0016787; F:hydrolyase activity; IEA.
 DR GO; GO:0004622; F:lysophospholipase activity; IEA.
 DR GO; GO:0009395; P:phospholipid catabolism; IEA.
 DR InterPro; IPR002642; PLAC.
 DR Pfam; PF01735; PLA2_B; 1.
 DR SMART; SM00022; PLAC; 1.
 DR Hydrolyase.
 SQ SEQUENCE 634 AA; 68281 MW; 657B72DEF08572D8 CRC64;
 QY
 Db 3 TGVLAGWGEKGKVFYPPDDKMSF 26
 164 TSLLENLWNIDSNLIFPDDKVSF 187

Query March 34.8%; Score 55; DB 3; Length 634;
 Best Local Similarity 37.5%; Pred. No. 29;
 Matches 9; Conservative 5; Mismatches 10; Indels 0; Gaps 0;

RESULT 8
 Q8S1Y5 PRELIMINARY; PRT; 195 AA.
 AC Q8S1Y5;
 DT 01-JUN-2002 (TREMREL. 21, Created)
 DT 01-OCT-2003 (TREMREL. 25, Last sequence update)
 DE Putative ubiquitin conjugating enzyme (EC 6.3.2.19) (Ubiquitin-
 DE conjugating enzyme E2) (Ubiquitin-protein ligase) (Ubiquitin carrier
 DE protein).
 GN P0482D04.9 OR OSUNBA0093F16.5.
 OS Oryza sativa (japonica cultivar-group).
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 OC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
 OC Ehrhartoideae; Oryzaceae; Oryza.
 OX NCBI_TaxID=39947;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=cv. Nipponbare;
 RA Sasaki T., Matsumoto T., Yamamoto K.;
 RT "Oryza sativa Nipponbare (GA3) genomic DNA, chromosome 1, PAC
 RT clone: P0482D04.";
 RL Submitted (FEB-2001) to the EMBL/GenBank/DBJ databases.
 RN [2]
 RP SEQUENCE FROM N.A.
 RC STRAIN=cv. Nipponbare;
 RA Sasaki T., Matsumoto T., Yamamoto K.;
 RT "Oryza sativa Nipponbare (GA3) genomic DNA, chromosome 1, PAC
 RT clone: P0482D04.";
 RL Submitted (FEB-2001) to the EMBL/GenBank/DBJ databases.
 CC -1- FUNCTION: CATALYZES THE COVALENT ATTACHMENT OF UBIQUITIN TO OTHER
 CC PROTEINS (BY SIMILARITY).
 CC -1- CATALYTIC ACTIVITY: ATP + UBIQUITIN + PROTEIN LYSINE = AMP +
 CC DIPOSPHATE + PROTEIN N-UBIQUITYLYSINE.
 CC -1- PATHWAY: UBIQUITIN CONJUGATION; SECOND STEP.
 CC -1- MISCELLANEOUS: A CYSTEINE RESIDUE IS REQUIRED FOR UBIQUITIN-
 CC THIOLESTER FORMATION (BY SIMILARITY).
 CC -1- SIMILARITY: BELONGS TO THE UBIQUITIN-CONJUGATING ENZYME FAMILY.
 DR EMBL; AP003262; BAB9862.1; -.
 DR EMBL; AP004332; BAB92885.1; -.
 DR Gramene; Q8S1Y5; -.
 DR GO; GO:0016874; F:ligase activity; IEA.
 DR GO; GO:0004640; F:ubiquitin conjugating enzyme activity; IEA.
 DR GO; GO:0004842; F:ubiquitin-protein ligase activity; IEA.

DR GO:0006512; P:ubiquitin cycle; IEA.
DR InterPro; IPR000449; UBA domain.
DR InterPro; IPR000608; Uba_conjugat.
DR Pfam; PF00627; UBA; 1.
DR Pfam; PF00179; UO_conj; 1.
DR Pfam; PF000461; Uba_conjugat; 1.
DR PROSITE; PS0030; UBA; 1.
DR PROSITE; PS00183; UBIQUITIN_CONJUGAT_1; 1.
DR PROSITE; PS0127; UBIQUITIN_CONJUGAT_2; 1.
DR Ligase; Ubi conjugation pathway.
SQ SEQUENCE 195 AA; 21323 MW; E78994BA038912B6 CRC64;

Query Match 34.2%; Score 54; DB 10; Length 195;
Best Local Similarity 43.2%; Pred. No. 11;
Matches 16; Conservative 3; Mismatches 6; Indels 12; Gaps 3;

Qy 3 TGVLAGWVG---EGGKFFV-----YFPD-DKMSFL 27
Db 37 TGTIAGPGCTPYEGSTFVIDRLPGYFPEPPKXQFI 73

RESULT 9

QBDTN6 PRELIMINARY; PRT; 310 AA.
AC QBDTN6;
DT 01-MAR-2003 (Tremblrel. 23, Created)
DT 01-MAR-2003 (Tremblrel. 23, Last sequence update)
DT 01-OCT-2003 (Tremblrel. 25, Last annotation update)
DE Conserved hypothetical protein.
GN SMU.1297.
OS Streptococcus mutans.
OC Bacteria; Firmicutes; Lactobacillales; Streptococcaceae;
OC Streptococcus.
CX NCBI_TaxID=1309;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=UA159 / ATCC 700610 / Serotype C;
RX MEDLINE=22295063; PubMed=12397186;
RA Ajdic D., Mchan W.M., McLaughlin R.E., Savic G., Chang J.,
RA Carson M.B., Primeaux C., Tian R., Kenton S., Jia H., Yin S., Qian Y.,
RA Li S., Zhu H., Najjar F., Lai H., White U., Roe B.A., Ferretti J.T.;
RT "Genome sequence of Streptococcus mutans UA159, a cariogenic dental
pathogen";
RL Proc. Natl. Acad. Sci. U.S.A. 99:14434-14439(2002).
DR EMBL; AE014964; AAN58974.1; -
DR GO:0003676; F:nucleic acid binding; IEA.
DR InterPro; IPR003156; DHAL.
DR InterPro; IPR001667; Psesterase.
DR Pfam; PF01368; DHH; 1.
DR Pfam; PF02272; DHAL; 1.
KW Hypothetical protein; Complete proteome.
SQ SEQUENCE 310 AA; 34625 MW; 17CE936D05E83587 CRC64;

Query Match 34.2%; Score 54; DB 16; Length 310;
Best Local Similarity 39.1%; Pred. No. 18;
Matches 9; Conservative 6; Mismatches 8; Indels 0; Gaps 0;

Qy 7 AGWGEKGKFFYPFDDKMSFLFA 29
Db 150 AGILGDTGRFLYPTATSKTFIA 172

RESULT 10

Q9BMG3 PRELIMINARY; PRT; 1087 AA.
AC Q9BMG3;
DT 01-JUN-2001 (Tremblrel. 17, Created)
DT 01-MAR-2002 (Tremblrel. 20, Last sequence update)
DT 01-OCT-2003 (Tremblrel. 25, Last annotation update)
DE Isoleucine tRNA synthetase (EC 6.1.1.5).
GN ILERS.
OS Leishmania donovani.
OC Eukaryota; Euklenozoa; Kinetoplastida; Trypanosomatidae; Leishmania.

OX NCBI_TaxID=5661;
RN [1]
RP SEQUENCE FROM N.A.
RA Abuduhani U., Selenke L., Crowe K., Mishra M., Lewis C.S.,
RA Chaudhuri G.;
RT "Cloning and characterization of the isoleucine tRNA synthetase gene
from Leishmania donovani";
RL Submitted (DEC-2001) to the EMBL/Genbank/DBJ databases.
DR EMBL; AF326935; AAG49529.2; -
DR HSSP; P56690; ILLE.
DR GO:0005524; P:ATP binding; IEA.
DR GO:0004822; P:Isoleucine-tRNA ligase activity; IEA.
DR GO:0016874; P:ligase activity; IEA.
DR GO:0006428; P:Isoleucyl-tRNA aminoacylation; IEA.
DR InterPro; IPR002300; tRNA-synt_1a.
DR InterPro; IPR001412; tRNA-synt_1a.
DR InterPro; IPR002301; tRNA-synt_1b.
DR InterPro; IPR009008; VALRS_ILERS_edit.
DR Pfam; PF00133; tRNA-synt_1_1.
DR PRINTS; PR00984; TRNASYNTHILS.
DR TIGRFAMs; TIGR00392; ILIS; 1.
DR PROSITE; PS00178; AA-tRNA_LIGASE_I; 1.
KW Aminoacyl-tRNA synthetase; ligase.
SQ SEQUENCE 1087 AA; 124462 MW; 11DD6AB2409146B8 CRC64;

Query Match 34.2%; Score 54; DB 5; Length 1087;
Best Local Similarity 50.0%; Pred. No. 72;
Matches 10; Conservative 4; Mismatches 6; Indels 0; Gaps 0;

Qy 7 AGWGEKGKFFYPFDDKMSF 26
Db 378 AGIFKGGKFFVCPVDENCMF 397

RESULT 11

Q7JEN7 PRELIMINARY; PRT; 294 AA.
AC Q7JEN7;
DT 01-OCT-2003 (Tremblrel. 25, Created)
DT 01-OCT-2003 (Tremblrel. 25, Last sequence update)
DT 01-OCT-2003 (Tremblrel. 25, Last annotation update)
DE Hypothetical protein.
GN RB1126.
OS Rhodospirillum rubrum.
OC Bacteria; Planctomycetes; Planctomycetacia; Planctomycetales;
OC Planctomycetaceae; Pirellula.
OX NCBI_TaxID=117;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=1;
RX MEDLINE=22735913; PubMed=12835416;
RA Gloeckner F.O., Kube M., Bauer M., Teeling H., Lombardot T.,
RA Ludwig W., Gade D., Beck A., Borzym K., Heltmann K., Rabus R.,
RA Schlesner H., Amann R., Reinhardt R.;
RT "Complete genome sequence of the marine planctomycete Pirellula sp.
strain 1";
RL Proc. Natl. Acad. Sci. U.S.A. 100:8298-8303(2003).
DR EMBL; BX294153; CAD78998.1; -
KW Hypothetical protein; Complete proteome.
SQ SEQUENCE 294 AA; 30660 MW; 55C8B8A1A78F8790 CRC64;

Query Match 33.5%; Score 53; DB 16; Length 294;
Best Local Similarity 58.8%; Pred. No. 24;
Matches 10; Conservative 4; Mismatches 3; Indels 0; Gaps 0;

Qy 1 VFTGLAGWGEKGKFFV 17
Db 183 LMTGVLSCGLPVGCGFV 199

RESULT 12

Q97016 PRELIMINARY; PRT; 854 AA.


```

AC Q97016;
DT 01-FEB-1997 (TrEMBLrel. 02, Created)
DT 01-FEB-1997 (TrEMBLrel. 02, Last sequence update)
DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
DE Envelope glycoprotein (Fragment).
GN ENV.
OS Human immunodeficiency virus 1.
OC Viruses; Retroviridae; Lentivirus.
OC NCBI_TaxID=11676;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=96190564; PubMed=8627686;
RA Gao F., Morrison S.G., Robertson D.L., Thornton C.L., Craig S.,
RA Karlsson G., Sodroski J., Morgado K., Galvao-Castro B.,
RA von Briesen H., Beddows S., Weber J., Sharp P.M., Shaw G.M.,
RA Hahn B.H.;
RT "Molecular cloning and analysis of functional envelope genes from
RT human immunodeficiency virus type 1 sequence subtypes A through G. The
RT WHO and NIAID Networks for HIV Isolation and Characterization.";
RI J. Virol. 70:1651-1657(1996).
RL [2]
RN [2]
RP SEQUENCE FROM N.A.
RA Allen E.E.;
RL Submitted (MAY-1995) to the EMBL/GenBank/DBJ databases.
DE EMBL: U27445; AAB06259.1; -.
DR GO: GO:0016021; C:Integral to membrane; IEA.
DR GO: GO:0019028; C:Viral capsid; IEA.
DR GO: GO:0019031; C:Viral envelope; IEA.
DR GO: GO:0005198; F:Structural molecule activity; IEA.
DR InterPro: IPR000328; Env GP41.
DR InterPro: IPR007777; GP120.
DR Pfam: PF00515; GP120; 1.
DR Pfam: PF00517; GP41; 1.
DR AIDS; Coat protein; Glycoprotein; Polypeptide; Transmembrane.
FT NON_TER
SQ SEQUENCE 854 AA; 96770 MW; 04A4B3196E2DD30D CRC64;

Query Match 33.5%; Score 53; DB 15; Length 854;
Best Local Similarity 57.1%; Pred. No. 77;
Matches 8; Conservative 2; Mismatches 4; Indels 0; Gaps 0;

Qy 6 LAGWGEKGKFPY 19
Db 583 LAGWGEKGKLIYP 596

RESULT 13
Q9J130 PRELIMINARY; PRT; 1151 AA.
AC Q9J130;
DT 01-OCT-2000 (TrEMBLrel. 15, Created)
DT 01-OCT-2000 (TrEMBLrel. 15, Last sequence update)
DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
DE Integrin beta 2 alpha subunit.
OC Rattus norvegicus (Rat).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
OC NCBI_TaxID=10116;
RN [1]
RP SEQUENCE FROM N.A.
RA Fathallah D.M. Sr., Zerrila K. Jr.;
RT "Cloning of the rat CD11b cDNA sequence.";
RL Submitted (MAY-2000) to the EMBL/GenBank/DBJ databases.
DE EMBL: AF268593; AAF81280.1; -.
DR HSSP: P11215; IBIQ.
DR GO: GO:0008305; C:Integrin complex; IEA.
DR GO: GO:0004885; F:cell adhesion receptor activity; IEA.
DR GO: GO:0007160; F:cell-matrix adhesion; IEA.
DR InterPro: IPR00413; Integrin_alpha.
DR InterPro: IPR002035; VWF_A.
DR Pfam: PF01839; FG-GAP; 3.
DR Pfam: PF00357; Integrin_A; 1.
DR Pfam: PF00092; vwa; 1.

```

```

DR PRINTS; PRO1185; INTEGRIN.
DR PRINTS; PRO0453; VWFADOMAIN.
DR SMART; SM00191; Int_alpha; 5.
DR SMART; SM00327; VWA; 1.
DR PROSITE; PS00242; INTEGRIN_ALPHA; 1.
DR PROSITE; PS00234; VWF_A; 1.
SQ SEQUENCE 1151 AA; 126943 MW; 8F765695D4074CA5 CRC64;

Query Match 33.5%; Score 53; DB 11; Length 1151;
Best Local Similarity 46.4%; Pred. No. 1.1e+02;
Matches 13; Conservative 2; Mismatches 9; Indels 4; Gaps 1;

Qy 4 GVLGVGE---GGKFPYPPDDKMSFL 27
Db 363 GVLGVSVSGFDWAGAFVPSKDKASFI 390

RESULT 14
Q9K1V4 PRELIMINARY; PRT; 4150 AA.
AC Q9K1V4;
DT 01-OCT-2000 (TrEMBLrel. 15, Created)
DT 01-OCT-2000 (TrEMBLrel. 15, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE 8,8a-deoxycleandiolide synthase 1.
GN CLEAT.
OS Streptomyces antibioticus.
OC Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;
OC Streptomyces; Streptomycetaceae; Streptomyces.
OC NCBI_TaxID=1890;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=20363406; PubMed=10908114;
RA Shah S., Xue Q., Tang L., Carney J.R., Betlach M., McDaniel R.;
RT "Cloning, Characterization, and Heterologous Expression of a
RT Polypeptide Synthase and P-450 Oxidase Involved in the Biosynthesis of
RT the Antibiotic Oleandomycin.";
RL J. Antibiot. 53:502-508(2000).
DE EMBL: AF220951; AAF82408.1; -.
DR HSSP: P25715; IMLA.
DR GO: GO:0004197; F:cysteine-type endopeptidase activity; IEA.
DR GO: GO:0016740; F:transferase activity; IEA.
DR GO: GO:0004314; F:lacyl-carrier protein] S-malonyltransferase. .; IEA.
DR GO: GO:0006633; P:fatty acid biosynthesis; IEA.
DR GO: GO:0008152; P:metabolism; IEA.
DR GO: GO:0006508; P:proteolysis and peptidolysis; IEA.
DR InterPro: IPR001227; Ac. trans.
DR InterPro: IPR004410; RabD.
DR InterPro: IPR000794; ketoacyl_synth.
DR InterPro: IPR006163; ppanthene_S.
DR InterPro: IPR006163; pp_bind.
DR InterPro: IPR000169; SHprot. acetyl.
DR Pfam: PF00698; Acyl_transf. 3.
DR Pfam: PF00109; ketoacyl-synt_C; 3.
DR Pfam: PF02801; ketoacyl-synt_C; 3.
DR Pfam: PF00550; pp-binding; 3.
DR TIGRPFAMs; TIGR00128; Iabd; 3.
DR PROSITE; PS00705; ACP_DOMAIN; 3.
DR PROSITE; PS00606; B_KETOACYL_SYNTHASE; 2.
DR PROSITE; PS00012; PHOSPHOPANTHETHEINE; 3.
DR PROSITE; PS00639; THIOL_PROTEASE_HIS; 1.
KV Phosphopantetheine; Transferase.
SQ SEQUENCE 4150 AA; 435261 MW; 9383296C4C16647D CRC64;

Query Match 33.5%; Score 53; DB 2; Length 4150;
Best Local Similarity 75.0%; Pred. No. 4.3e+02;
Matches 9; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

Qy 7 AGWGEKGKFPY 18
Db 3838 AGWGEKGKQAVY 3849

```

RESULT 15

```

QSRD34      QSRD34      PRELIMINARY;      PRT;      355 AA.
ID          QSRD34
AC          QSRD34;
DT          01-JUN-2002 (TRENBLREL. 21, Created)
DT          01-JUN-2002 (TRENBLREL. 21, Last sequence update)
DT          01-MAR-2003 (TRENBLREL. 23, Last annotation update)
DE          Beta-lactamase class C and other penicillin binding proteins.
GN          AMPC OR TTE0213.
OS          Thermomanaerobacter tengcongensis.
OC          Bacteria; Firmicutes; Clostridia; Thermomanaerobacteriales;
OC          Thermomanaerobacteriaceae; Thermomanaerobacter.
OX          NCBI_TaxID=119072;
RN          [1]
RP          SEQUENCE FROM N.A.
RC          STRAIN=ME4 / JCM 11007;
RX          MEDLINE=2192816; PubMed=1197336;
RA          Bao Q., Tian Y., Li W., Xu Z., Xuan Z., Hu S., Dong W., Yang J.,
RA          Chen Y., Xue Y., Xu Y., Lai X., Huang L., Dong X., Ma Y., Ling L.,
RA          Tan H., Chen R., Wang J., Yu J., Yang H.;
RT          "A complete sequence of T. tengcongensis genome.";
RL          Genome Res. 12:689-700(2002).
DR          EMBL; AE012994; AAM23514.1; -
DR          InterPro; IPR001466; Beta_lactamase.
DR          Pfam; PF00144; beta-lactamase; 1.
KW          Complete Proteome.
SQ          SEQUENCE 355 AA; 40303 MW; 3FD7B0ADAE46545 CRC64;

Query Match          33.2%; Score 52.5; DB 16; Length 355;
Best Local Similarity 35.3%; Pred. No. 35;
Matches 12; Conservative 6; Mismatches 7; Indels 9; Gaps 1;

QY          1 VETGVLAGVWGE-----GGRKYYPPDDKXS 25
              |||||
DB          23 VFGVAAAAGDDKGFVRVEVKNRKLXPDEKLN 56

```

Search completed: September 30, 2004, 11:08:32
 Job time : 63.0381 secs


```

RESULT 2
HEPA_HSV2H
ID_HEPA_HSV2H STANDARD; PRT; 752 AA.
AC P89431;
DT 01-NOV-1997 (Rel. 35, Created)
DT 01-NOV-1997 (Rel. 35, Last sequence update)
DT 01-NOV-1997 (Rel. 35, Last annotation update)
DE DNA helicase/primase complex associated protein.
GN UL8.
OS Herpes simplex virus (type 2 / strain HG52).
CC Viruses; dsDNA viruses, no RNA stage; Herpesviridae;
CC Alphaherpesvirinae; Simplexvirus.
OX NCBI_TaxID=10315;
RN [1]
RP SEQUENCE FROM N.A.
RA Dolan A.;
RL Submitted (FEB-1997) to the EMBL/GenBank/DBJ databases.
CC -1- FUNCTION: PROBABLY INVOLVED IN DNA REPLICATION.
CC -1- SIMILARITY: BELONGS TO FAMILY THAT GROUPS TOGETHER HSV-1 UL8,
CC EHV-1 54, VZV 52 AND HCMV 102.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; Z86099; CAB06768.1; -
DR InterPro; IPR004996; Herpes_HEPA.
DR Pfam; PF03324; Herpes_HEPA.1.
KW DNA replication.
SQ SEQUENCE 752 AA; 80026 MW; 8749B92360B58AD5 CRC64;

Query Match
Best Local Similarity 82.9%; Score 131; DB 1; Length 752;
Pred. No. 5,7e-11;
Matches 23; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

QY 1 VFTGVLAGWGEKGKFFYPDDKMSFLP 28
DB 724 VFTGVLEGWGEKGSFYFPEDEKIRFLP 751

RESULT 3
Cl56_STRAT STANDARD; PRT; 3519 AA.
AC Q07017;
DT 01-NOV-1997 (Rel. 35, Created)
DT 01-NOV-1997 (Rel. 35, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Oleandomycin polyketide synthase, modules 5 and 6.
GN ORB5.
OS Streptomyces antibioticus.
CC Bacteria; Actinobacteriae; Actinobacteridae; Actinomycetales;
CC Streptomycinae; Streptomycetaceae; Streptomyces.
OX NCBI_TaxID=1890;
RN [1]
RP SEQUENCE FROM N.A.
RA MEDLINE=94150470; PubMed=8107683;
RA Swan D.G.; Rodriguez A.M.; Vilches C.; Mendez C.; Salas J.A.;
RT "Characterisation of a Streptomyces antibioticus gene encoding a type
RT I polyketide synthase which has an unusual coding sequence.";
RL Mol. Genet. 242:358-362(1994).
CC -1- FUNCTION: MAY BE INVOLVED IN THE BIOSYNTHESIS OF THE OLEANDOMYCIN
CC LACTONE RING.
CC -1- COFACTOR: Contains 2 covalently bound phosphopantetheines.
CC -1- SIMILARITY: Contains 2 acyl carrier domains.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -

```

```

CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; L09654; AAA19659.1; -
DR PIR; S43048; S43048.
DR HSSP; F25715; IMUA.
DR InterPro; IPR001277; Ac_trans.
DR InterPro; IPR000794; Ketoacyl_synth.
DR InterPro; IPR006163; Pp_bind.
DR InterPro; IPR006162; Panine_S.
DR InterPro; IPR001031; Thioesterase.
DR Pfam; PF00698; Acyl_transf. 2.
DR Pfam; PF0109; Ketoacyl-synt. 2.
DR Pfam; PF02801; Ketoacyl-synt_C; 2.
DR Pfam; PF00550; Pp-binding; 2.
DR Pfam; PF00975; Thioesterase; 1.
DR PROSITE; PS00012; PHOSPHOPANTHETHEINE; 2.
DR PROSITE; PS00606; B_KETOACYL_SYNTHASE; 2.
DR PROSITE; PS50975; ACP_DOMAIN; 2.
KW Transferase; Acyltransferase; Antibiotic biosynthesis; NADP;
KW Phosphopantetheine; Multifunctional enzyme; Repeat.
FT DOMAIN 1 3519
FT DOMAIN 2 32 501
FT DOMAIN 3 569 890
FT DOMAIN 4 1280 1382
FT DOMAIN 5 1467 1561
FT DOMAIN 6 1686 2156
FT DOMAIN 7 2220 2541
FT DOMAIN 8 2856 3038
FT DOMAIN 9 3141 3215
FT DOMAIN 10 3270 3519
FT ACT_SITE 210 210
FT ACT_SITE 660 660
FT NP_BIND 1203 1249
FT BINDING 1524 1524
FT ACT_SITE 1859 1859
FT ACT_SITE 2311 2311
FT NP_BIND 2859 2905
FT BINDING 3178 3178
SQ SEQUENCE 3519 AA; 368561 MW; 41AE78AAAB61F86 CRC64;

Query Match
Best Local Similarity 33.5%; Score 53; DB 1; Length 3519;
Pred. No. 37;
Matches 9; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 7 AGWGEKGKFFY 18
DB 1337 AGWGSQGQAVY 1348

RESULT 4
DCTA_BACSU
ID_DCTA_BACSU STANDARD; PRT; 421 AA.
AC P96503;
DT 28-FEB-2003 (Rel. 41, Created)
DT 28-FEB-2003 (Rel. 41, Last sequence update)
DT 15-MAR-2004 (Rel. 43, Last annotation update)
DE C4-dicarboxylate transport protein (C4-dicarboxylate permease).
GN DCTA OR DCTP OR BSU04470.
OS Bacillus subtilis.
CC Bacteria; Firmicutes; Bacillales; Bacillaceae; Bacillus.
OX NCBI_TaxID=1423;
RN [1]
RP SEQUENCE FROM N.A.
RA STRAIN=168;
RA Kaashara Y.; Nakai S.; Sadaie Y.; Ogasawara N.;
RT "A 148 kbp sequence of the region between 35 and 47 degree of the
RT Bacillus subtilis genome.";
RL Submitted (MAR-1997) to the EMBL/GenBank/DBJ databases.

```

[2] SEQUENCE FROM N.A.

RP STRAIN=168;
RX MEDLINE=98044033; PubMed=9384377;
RA Kunst F., Ogatawara N., Moszer I., Albertini A.M., Alloni G.,
RA Azevedo V., Bertero M.G., Bessieres P., Bolotin A., Borchert S.,
RA Borries R., Boulter L., Brans A., Braun M., Briggell S.C., Bron S.,
RA Brouillet S., Brusch C.V., Caldwell I.B., Capuano V., Carter N.M.,
RA Choi S.K., Codani J.J., Connerton I.F., Cummings N.J., Daniel R.A.,
RA Denizot F., Devine K.M., Dusterhoft A., Enrich S.D., Emerson P.T.,
RA Entian K.D., Ertington J., Fabre C., Ferrati E., Foulger D.,
RA Fritz C., Fujita M., Fujita Y., Fuma S., Galizzi A., Galleron N.,
RA Gilm S.Y., Glaser P., Goffeau A., Golightly E.J., Grand G.,
RA Guisepi G., Guy B.J., Haga K., Hachez J., Harwood C.R., Hentut A.,
RA Hilbert H., Holtsappel S., Hosono S., Hullio M.F., Itaya M., Jones L.,
RA Joris B., Karamata D., Kashara Y., Kjaer-blanchard M., Klein C.,
RA Kodayaehi Y., Koester P., Koningsseim G., Krogh S., Kumano M.,
RA Kutita K., Lapidis A., Lardinois S., Laber J., Lazarevic V.,
RA Lee S.M., Levine A., Liu H., Masuda S., Maul C., Medgye C.,
RA Medina N., Mellado R.P., Mizuno M., Moestl D., Nakai S., Nockack M.,
RA Noone D., O'Reilly M., Ogawa K., Ogihara A., Oudega B., Park S.H.,
RA Parro V., Pohl T.M., Portelle D., Porollik S., Prescott A.M.,
RA Pressac E., Puig C., Purnelle B., Rapoport G., Rey M., Reynolds S.,
RA Rieger N., Rivolta C., Rocca E., Roche B., Rose M., Sadate Y.,
RA Sato T., Scanlan E., Schleich S., Schroeder R., Scofield F.,
RA Sekiguchi Y., Sekowska A., Sedor S.U., Serio P., Shin B.S., Soldo B.,
RA Sorokin A., Taccoti E., Takagi T., Takahashi H., Takenaru K.,
RA Takeuchi M., Tamakoshi A., Tanaka T., Terstera P., Tognoni A.,
RA Tosato V., Uchiyama S., Vanderbol M., Vannier F., Vassarotti A.,
RA Viari A., Wandut R., Wedler E., Weiler H., Weitzneger T.,
RA Winters P., Wipac A., Yamamoto H., Yamane K., Yasunaga K., Yata K.,
RA Yoshida K., Yoshikawa H.F., Zumsstein E., Yoshikawa H., Danchin A.,
RT "The complete genome sequence of the Gram-positive bacterium *Bacillus*
RL *subtilis*";
RT Nature 390:245-256(1997).

[3] FUNCTION, AND GENE NAME.

RP STRAIN=168;
RX MEDLINE=20170658; PubMed=10708364;
RA Asai K., Baik S.-H., Kasahara Y., Moriya S., Ogatawara N.,
RT "Regulation of the transport system for C4-dicarboxylic acids in
RL *Bacillus subtilis*";
RL Microbiology 146:263-271(2000).

-1- FUNCTION: Responsible for the transport of succinate and fumarate,
CC but not malate, across the membrane (Probable).
CC -1- SUBCELLULAR LOCATION: Integral membrane protein (Probable).
CC -1- INDUCTION: Induced by succinate and fumarate. Expression
CC of DctA requires the DctB, DctS and DctP proteins and is decreased
CC in the presence of malate.
CC -1- SIMILARITY: Belongs to the sodium:dicarboxylate (SDF) symporter
CC (TC 2.A.23) family.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by, and for commercial
CC entities requires a license agreement (see <http://www.isb-sib.ch/announce/>).
CC or send an email to license@isb-sib.ch.

CC EMBL; AB001488; BAA19284.1; -
DR EMBL; Z99106; CAB12554.1; -
DR PIR; G69771; G69771.
DR Subtilist; BG12075; dctA.
DR HAMAP; MF_01300; -; 1.
DR InterPro; IPR001991; Na/dico_symport.
DR Pfam; PF00375; SDF; 1.
DR PRINTS; PR00173; EDPNRSPT.
DR PROSITE; PS00713; NA_DICARBOXYL_SYMPT_1; FALSE_NEG.
DR PROSITE; PS00714; NA_DICARBOXYL_SYMPT_2; 1.
KM Transport; Sugar transport; Transmembrane; Symport; Complete proteome.
FT TRANSMEM 9 29
FT TRANSMEM 39 59
FT POTENTIAL.

	FT	TRANSMEM	76	96	POTENTIAL.
	FT	TRANSEM	145	165	POTENTIAL.
	FT	TRANSMEW	185	205	POTENTIAL.
	FT	TRANSEWE	219	239	POTENTIAL.
	FT	TRANSEM	316	336	POTENTIAL.
	FT	TRANSEM	348	368	POTENTIAL.
SQ	SEQUENCE		421 AA;	45441 MW; 29FA5A6BD5DD7A5A CRC64;	
Query March			32.3%; Score 51; DB 1;	Length 421;	
Best Local Similarity			43.8%;	Pred No 8.3;	Mismatches 4; Gaps 1;
Matches 14;	Conservative		4;	Mismatches 10;	Gaps 4;
Dn	151 LFFSILFGVGHAAIGEKSGKSVIDPFDKXSHVF	182			
QY	1 VETGVLAVGV---WGEGGRFYVPFDDYMSLFLP	28			
	: :	:			
	I F T I A S Y E R D P H Q M K N W C S	: :			
	L I I I I I I I I I I I I I I I I I I	:			
CAS CANFA	ID_CXAS_CANFA	STANDARD,	PRT;	356 AA.	
AC	P33725;				
DT	01-FEB-1994 (Rel., 28, Created)				
DT	01-FEB-1994 (Rel., 28, Last sequence update)				
DT	28-FEB-2003 (Rel., 41, Last annotation update)				
DE	Gap junction alpha-5 protein (Connexin 40) (Cx40).				
GN	GU.S.				
OS	Canis familiaris (dog).				
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;				
OC	Mammalia; Eutheria; Carnivora; Fissipedia; Canidae; Canis.				
OX	NCBI_Taxid=9615;				
RN	[1]				
RE	SEQUENCE FROM N.A.				
RX	MEDLINE=92136511; PubMed=1310450;				
RA	Kanter B.U., Safitz U.E., Beyer E.C.;				
"Cardiac myocytes express multiple gap junction proteins."					
Circ. Res. 70:438-444(1992).					
-I- FUNCTION: One gap junction consists of a cluster of closely packed pairs of transmembrane channels, the connexons, through which materials of low MW diffuse from one cell to a neighboring cell.					
-I- SUBUNIT: A connexon is composed of a hexamer of connexins.					
-I- SUBCELLULAR LOCATION: Integral membrane protein.					
-I- SIMILARITY: Belongs to the connexin family. Alpha-type (group II) subfamily.					
--					
CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL Outstation - The European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See http://www.isb-sib.ch/announce/isb-sib.ch).					
OR send an email to license@lsb-sib.ch).					
DR EMBL; M81347; AAA30838.1; -. .					
PTM; A49024; A49024.					
DR InterPro; IPRO00500; Connexin.					
DR Pfam; PF00029; connexin_1.					
DR PRINTS; PR00206; CONNXIN.					
DR SMART; SMO0037; CNX; 1.					
DR PROSITE; PS00407; CONNEXXINS_1; 1.					
DR PROSITE; PS00408; CONNEXXINS_2; 1.					
KW Gap Junction; Transmembrane.					
FT DOVAIN	0	BY SIMILARITY			
FT DOVAIN	1	CYTOLASMIC (POTENTIAL).			
FT TRANSMEW	23	POTENTIAL.			
FT DOVAIN	41	EXTRACELLULAIR (POTENTIAL).			
FT TRANSMEW	77	POTENTIAL.			
FT DOVAIN	100	CYTOLASMIC (POTENTIAL).			
FT TRANSMEW	155	POTENTIAL.			
FT DOVAIN	178	EXTRACELLULAIR (POTENTIAL).			
FT TRANSMEW	205	POTENTIAL.			
FT DOVAIN	228	CYTOLASMIC (POTENTIAL).			
Q SEQUENCE	356 AA; 39801 MW; C8431D65CFE578FE CRC64;				

```

Query Match      31.6%; Score 50; DB 1; Length 356;
Best Local Similarity 60.0%; Pred. NO. 9.8;
Matches 9; Conservative 2; Mismatches 4; Indels 0; Gaps 0;

QY 11 GEGGKFFVPPDDKMS 25
Db 272 GPGGKFFNPFNSKMA 286

RESULT 6
YCDZ_SALTY STANDARD; PRT; 160 AA.
ID YCDZ_SALTY STANDARD; PRT; 160 AA.
AC 054290;
DT 15-DEC-1998 (Rel. 37, Created)
DT 28-FEB-2003 (Rel. 41, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Hypothetical protein ycdz.
GN YCDZ OR STM138.
OS Salmonella typhimurium.
OC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
OC Enterobacteriaceae; Salmonella.
OC NCBI_Taxid=602;
RN SEQUENCE FROM N.A.
RC STRAIN=LT2 / SGC1412 / ATCC 700720;
RX MEDLINE=21534948; PubMed=11677609;
RA McClelland M., Sanderson K.E., Spieth J., Clifton S.W., Latreille P.,
RA Courtney L., Porwollik S., Ali J., Dante M., Du P., Hou S., Layman D.,
RA Leonard S., Nguyen C., Scott K., Holmes A., Grewal N., Mulvaney E.,
RA Ryan E., Sun H., Flores L., Miller W., Storey T., Nhan M.,
RA Waterston R., Wilson R.K.;
RT "Complete genome sequence of Salmonella enterica serovar Typhimurium
RT LT2."
RL Nature 413:852-856(2001).
RN SEQUENCE OF 55-160 FROM N.A.
RC STRAIN=SR-11;
RX MEDLINE=98117058; PubMed=9457880;
RA Romling U., Bian Z., Hammar M., Sierraalta W.D., Normark S.;
RT "Curli fibers are highly conserved between Salmonella typhimurium and
RT Escherichia coli with respect to operon structure and regulation."
RL Bacteriol. 180:722-731(1998).
RN SUBCELLULAR LOCATION: Integral membrane protein (Potential).
CC -1- SIMILARITY: TO E.COLI YACB.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; AE008749; AAJ20068.1; ALT_INIT.
DR EMBL; AJ002301; CA05311.1; -.
DR StyGene; SG10739; ycdz.
KW Hypothetical protein; Transmembrane; Complete proteome.
FT TRANSMEM 20 42 POTENTIAL.
FT TRANSMEM 50 70 POTENTIAL.
FT TRANSMEM 72 92 POTENTIAL.
FT TRANSMEM 99 119 POTENTIAL.
FT TRANSMEM 123 143 POTENTIAL.
SQ SEQUENCE 160 AA; 16619 MW; SE10D7EDCEBA7DA CRC64;

Query Match      31.0%; Score 49; DB 1; Length 160;
Best Local Similarity 53.8%; Pred. NO. 6;
Matches 7; Conservative 4; Mismatches 2; Indels 0; Gaps 0;

QY 1 VFTGLAGVWGEG 13
Db 9 ITTGILSGIMWG 21

```

```

RESULT 7
ID SYN_YEAST STANDARD; PRT; 546 AA.
AC P07263; Q12095;
DT 01-APR-1988 (Rel. 07, Created)
DT 01-NOV-1997 (Rel. 35, Last sequence update)
DT 10-OCT-2003 (Rel. 42, Last annotation update)
DE Histidyl-tRNA synthetase, mitochondrial precursor (EC 6.1.1.21)
DE (Histidine--tRNA ligase) (HISRS).
GN HTS1 OR YPR033C OR YP9367.13C.
OS Saccharomyces cerevisiae (Baker's yeast).
OC Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
OC Saccharomycetales; Saccharomycetaceae; Saccharomyces.
OC NCBI_Taxid=4932;
RN SEQUENCE FROM N.A.
RP MEDLINE=86245069; PubMed=3521891;
RX Natsoulis G., Hilger F., Fink G.R.;
RT "The HTS1 gene encodes both the cytoplasmic and mitochondrial
RT histidine tRNA synthetases of S. cerevisiae."
RL Cell 46:235-243(1986).
RN SEQUENCE FROM N.A.
RP STRAIN=S288c / AB972;
RX MEDLINE=97313271; PubMed=9169875;
RA Bussey H., Storms R.K., Ahmed A., Albermann K., Allen E., Ansgore W.,
RA Araujo R., Aparicio A., Bartell B.G., Badcock K., Benes V.,
RA Bortwein D., Bowman S., Bruckner M., Carpenter J., Cherry J.M.,
RA Chung E., Churcher C.M., Coster F., Davis K., Davis R.W.,
RA Duran F.S., Delius H., Dipolo T., Dubois E., Duesterhoeft A.,
RA Duncan M., Floeth M., Fortin N., Friesen J.D., Fritz C., Goffeau A.,
RA Hall J., Hebling U., Heumann K., Hilbert H., Hillier L.W.,
RA Huntke-Smith S., Hyman R., Johnston W., Kaiman S., Kleene K.,
RA Komp C., Kurd O., Lashkari D., Lew H., Lin A., Lin D., Louis E.J.,
RA Marathe R., Messenguy F., Mewes H.-W., Mitrpali S., Moestl D.,
RA Mueller-Auer S., Namath A., Newnich U., Oefner P., Pearson D.,
RA Petel F.X., Pohl T.W., Purnelle D., Schaefer M., Scharfe M.,
RA Scherens B., Schamm S., Schroeder M., Sedou A.M., Tettein H.,
RA Uristarazu L.A., Ushinsky S., Verendeels F., Vissers S., Voss H.,
RA Walsh S.V., Wandt R., Wang Y., Wedler E., Wedler H., Winnett E.,
RA Zhong W.W., Zollner A., Vo D.H., Han J.;
RT "The nucleotide sequence of Saccharomyces cerevisiae chromosome XVI."
RL Nature 387:103-105(1997).
RN [3]
RN MUTAGENESIS.
RX MEDLINE=93093458; PubMed=1459448;
RA Cain M.I., Mason T.L., Fink G.R.;
RT "HTS1 encodes both the cytoplasmic and mitochondrial histidyl-tRNA
RT synthetase of Saccharomyces cerevisiae: mutations alter the
RT specificity of compartmentation."
RL Genetics 132:987-1001(1992).
CC -1- CATALYTIC ACTIVITY: ATP + L-histidine + tRNA(His) = AMP +
CC diphosphate + L-histidyl-tRNA(His).
CC -1- SUBCELLULAR LOCATION: Mitochondrial and cytoplasmic.
CC -1- ALTERNATIVE PRODUCTS:
CC Event=Alternative initiation;
CC Comment=2 isoforms, Mitochondrial (shown here) and Cytoplasmic,
CC are produced by alternative initiation;
CC -1- SIMILARITY: Belongs to class-II aminoacyl-tRNA synthetase family.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; M14048; AAA34695.1; -.
DR EMBL; M14048; AAA34696.1; -.
DR EMBL; Z71255; CAA94983.1; -.
DR EMBL; Z49274; CAA89287.1; -.
DR PIR; S54507; SYBYHM.

```

```

DR Germonline; 144298; -.
DR SGD; S0006237; HTSL.
DR GO; GO:0005739; C:mitochondrion; IMP.
DR GO; GO:0006427; P:histidyl-trna aminoacylation; IMP.
DR InterPro: IPR004154; HGRF-anticodon.
DR InterPro: IPR004516; Hiss.
DR InterPro: IPR003314; tRNA-synt-2b.
DR InterPro: IPR006195; tRNA-ligase-II.
DR Pfam; PF03129; HGRF-anticodon; 1.
DR Pfam; PF00587; tRNA-synt-2b; 1.
DR TIGRFAMs; TIGR00442; hss; 1.
DR PROSITE; PS50862; AA TRNA-LIGASE II; 1.
KW Aminoacyl-trna synthetase; Protein biosynthesis; Ligase; ATP-binding;
KM Mitochondrion; Transic peptide; Alternative initiation.
FT TRANSIT 1 20 MITOCHONDRION.
FT CHAIN 21 546 HISTIDYL-TRNA SYNTHETASE, ISOFORM
FT CHAIN 22 546 MITOCHONDRIAL.
FT CHAIN 22 546 HISTIDYL-TRNA SYNTHETASE, ISOFORM
FT INIT MET 21 21 FOR ISOFORM CYTOPLASMIC.
FT CONFLICT 476 478 AAE -> TTK (IN RRF. 1).
SQ SEQUENCE 546 AA; 59952 MW; 8BDCCB4BD73973 CRC64;

Query Match 31.0%; Score 49; DB 1; Length 546;
Best Local Similarity 42.1%; Pred. No. 21;
Matches 8; Conservative 5; Mismatches 6; Indels 0; Gaps 0;

QY 5 VLAGWGEGGKFFYPDDK 23
Db 101 ILAGKYGEDSKLYNLEDQ 119

RESULT 8
MUTY_HAEIN STANDARD; PRT; 378 AA.
ID MUTY_HAEIN
AC P44320;
DT 01-NOV-1995 (Rel. 32, Created)
DT 01-NOV-1995 (Rel. 32, Last sequence update)
DT 26-FEB-2003 (Rel. 41, Last annotation update)
DE A/G-specific adenine glycosylase (EC 3.2.2.-).
GN MUTY OR H10759.
OS Haemophilus influenzae.
OC Bacteria; Proteobacteria; Gammaproteobacteria; Pasteurellales;
OC Pasteurellaceae; Haemophilus.
OX NCBI_TaxID=727;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=Rd / KW20 / ATCC 51907;
RX MEDLINE=95350630; PubMed=7542800;
RA Fleischmann R.D., Adams M.D., White O., Clayton R.A., Kirkness E.F.,
RA Kerlavage A.R., Bult C.J., Tomb J.-F., Dougherty B.A., Merrick J.M.,
RA McInerney K., Sutton G., Fitzhugh W., Fields C.A., Gocayne J.D.,
RA Scott J.D., Shiley R., Liu L.-I., Glodek A., Kelley J.M.,
RA Weidman J.F., Phillips C.A., Spriggs T., Hedblom E., Cotton M.D.,
RA Uterback T.R., Hanna M.C., Nguyen D.T., Saudek D.M., Brandon R.C.,
RA Gnehm C.L., McDonald L.A., Small K.V., Fraser C.M., Smith H.O.,
RA Venter J.C.;
RT "Whole-genome random sequencing and assembly of Haemophilus influenzae
RT Rd.";
RL Science 269:496-512(1995).
CC -1- FUNCTION: Adenine glycosylase active on G-A and C-A mispairs (By
CC similarity).
CC -1- COFACTOR: Binds a 4Fe-4S cluster which is not important for the
CC catalytic activity, but which is probably involved in the proper
CC positioning of the enzyme along the DNA strand (By similarity).
CC -1- SIMILARITY: Belongs to the nhb/mutY family.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial

```

```

CC entities requires a license agreement (see http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; U32760; AAC22418.1; -.
DR PIR; C64091; C64091.
DR HSSP; P17802; 1MOV.
DR TIGR; H10759; -.
DR InterPro: IPR003265; Endo-3c.
DR InterPro: IPR004035; EndoIII_PCL.
DR InterPro: IPR004036; EndoIII_HNH.
DR InterPro: IPR003651; Hss bind.
DR InterPro: IPR004445; Hnh.
DR InterPro: IPR005760; MutY.
DR Pfam; PF00730; HNH-GPD; 1.
DR Pfam; PF00633; HNH; 1.
DR SMART; SM00478; ENDO3C; 1.
DR SMART; SM00525; FES; 1.
DR TIGRFAMs; TIGR01084; mutY; 1.
DR PROSITE; PS00764; ENDONUCLEASE_III_1; 1.
DR PROSITE; PS01155; ENDONUCLEASE_III_2; 1.
KW DNA repair; Hydrolyase; Glycosylase; Iron-sulfur; 4Fe-4S;
KM Complete proteome.
FT METAL 197 IRON-SULFUR (4FE-4S) (BY SIMILARITY).
FT METAL 204 IRON-SULFUR (4FE-4S) (BY SIMILARITY).
FT METAL 207 IRON-SULFUR (4FE-4S) (BY SIMILARITY).
FT METAL 213 IRON-SULFUR (4FE-4S) (BY SIMILARITY).
SQ SEQUENCE 378 AA; 43505 MW; C443FE25131B2A21 CRC64;

Query Match 30.7%; Score 48.5; DB 1; Length 378;
Best Local Similarity 50.0%; Pred. No. 17;
Matches 11; Conservative 4; Mismatches 4; Indels 3; Gaps 2;

QY 7 AGVWGEGGKFFYP-FDDXMSFL 27
Db 257 SGLM--GGLFCRFQFEDKXSL 276

RESULT 9
YA17_HAEIN STANDARD; PRT; 213 AA.
ID YA17_HAEIN
AC O86231;
DT 30-MAY-2000 (Rel. 39, Created)
DT 30-MAY-2000 (Rel. 39, Last sequence update)
DT 10-OCT-2003 (Rel. 42, Last annotation update)
DE Hypothetical aquaporin-like protein H1017.
GN H1017.
OS Haemophilus influenzae.
OC Bacteria; Proteobacteria; Gammaproteobacteria; Pasteurellales;
OC Pasteurellaceae; Haemophilus.
OX NCBI_TaxID=727;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=Rd / KW20 / ATCC 51907;
RX MEDLINE=95350630; PubMed=7542800;
RA Fleischmann R.D., Adams M.D., White O., Clayton R.A., Kirkness E.F.,
RA Kerlavage A.R., Bult C.J., Tomb J.-F., Dougherty B.A., Merrick J.M.,
RA McInerney K., Sutton G., Fitzhugh W., Fields C.A., Gocayne J.D.,
RA Scott J.D., Shiley R., Liu L.-I., Glodek A., Kelley J.M.,
RA Weidman J.F., Phillips C.A., Spriggs T., Hedblom E., Cotton M.D.,
RA Uterback T.R., Hanna M.C., Nguyen D.T., Saudek D.M., Brandon R.C.,
RA Gnehm C.L., Fritchman J.L., Fuhrmann J.L., Geoghegan N.S.M.,
RA Pine L.D., McDonald L.A., Small K.V., Fraser C.M., Smith H.O.,
RA Venter J.C.;
RT "Whole-genome random sequencing and assembly of Haemophilus influenzae
RT Rd.";
RL Science 269:496-512(1995).
CC -1- IDENTIFICATION.
CC -----
CC White O., Clayton R.A., Kerlavage A.R., Fleischmann R.D., Peterson J.,
CC Hickey B., Dodson R., Gwinn M.,
CC Submitted (MAY-1998) to the EMBL/GenBank/DBJ databases.
CC -1- SUBCELLULAR LOCATION: Integral membrane protein (potential).
CC -----
CC -1- DOMAIN: Aquaporins contain two tandem repeats each containing

```


RESULT 14
Y087_CAEEL STANDARD; PRT; 324 AA.
ID Y087_CAEEL
AC P34623; Q05035;
DT 01-FEB-1994 (Rel. 28, Created)
DT 28-FEB-2003 (Rel. 41, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Hypothetical protein ZK1236.7 in chromosome III.
GN ZK1236.7/C30011.3
OS Caenorhabditis elegans.
OC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditoidea;
OC Rhabditidae; Pelodierinae; Caenorhabditis.
OX NCBI_TaxID=6239;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=Bristol N2;
RX MEDLINE=94150718; PubMed=7906398;
RA Wilson R., Anscough R., Anderson K., Baynes C., Berks M.,
RA Bonfield J., Burton J., Connell M., Copsey T., Cooper J., Coulson A.,
RA Cretton M., Dear S., Du Z., Durbin R., Favell A., Fraser A.,
RA Fulton L., Gardner A., Green P., Hawkins T., Hillier L., Jier M.,
RA Johnson L., Jones M., Kershaw J., Kirsten J., Laister N.,
RA Larrelle P., Lightning J., Lloyd C., Mortimore B., O'Callaghan M.,
RA Parsons J., Percy C., Rifkin L., Roopra A., Saunders D., Shearman R.,
RA Sims M., Smalton N., Smith A., Smith M., Sonhammer G., Staden R.,
RA Sutcliffe J., Thierry-Mieg J., Thomas K., Vaubin M., Vaughan K.,
RA Watson R., Watson A., Weinstock L., Wilkinson-Spratt J.,
RA Wohldman P.;
RT "2.2 Mb of contiguous nucleotide sequence from chromosome III of C.
RT elegans";
RL Nature 368:32-38 (1994).
RN [2]
RP REVISIONS.
RA Waterston R.;
RL Submitted (OCT-2001) to the EMBL/GenBank/DBJ databases.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL: L13200; AAA28188.2; -.
DR PIR: S44785; S44785.
DR WormPep; ZK1236.7; CE28628.
DR InterPro; IPR000717; PCI.
KW Hypothetical protein.
FT DOMAIN 143 226 GLU-RICH.
SQ SEQUENCE 324 AA; 37446 MW; AED779E03804435D CRC64;
Query Match 29.7%; Score 47; DB 1; Length 324;
Best Local Similarity 36.4%; Pred. No. 24;
Matches 8; Conservative 7; Mismatches 7; Indels 0; Gaps 0;
CY 4 GVLAVGWGEGKGFVPPDDKMS 25
Db 269 GLVGVWMDRGKRFYISDEEFA 290
RESULT 15
LACY_LEULA STANDARD; PRT; 639 AA.
ID LACY_LEULA
AC Q48624;
DT 15-DEC-1998 (Rel. 37, Created)
DT 15-DEC-1998 (Rel. 37, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Lactose permease (lactose-proton symport) (Lactose transport protein).
GN LACS.
OS Leuconostoc lactis.

OG Plasmid pN263.
OC Bacteria; Firmicutes; Lactobacillales; Leuconostoc.
OX NCBI_TaxID=1246;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=N26009;
RX MEDLINE=96209221; PubMed=8633855;
RA Vaughan E.E., David S., de Vos W.M.;
RT "The lactose transporter in Leuconostoc lactis is a new member of the
RT lacs subfamily of galactoside-pentose-hexuronide translocators";
RL Appl. Environ. Microbiol. 62:1574-1582 (1996).
CC -1- FUNCTION: RESPONSIBLE FOR TRANSPORT OF BETA-GALACTOSIDES INTO THE
CC CELL, WITH THE CONCOMITANT EXPORT OF A PROTON (SYMPORT SYSTEM) OF
CC AND ALSO FOR TRANSPORT OF HOMOLOGOUS AND HETEROLOGOUS EXCHANGE OF
CC BETA-GALACTOSIDES.
CC -1- SUBCELLULAR LOCATION: Integral membrane protein.
CC -1- DOMAIN: THE ENZYME IIA-LIKE REGION MAY SERVE A REGULATORY
CC FUNCTION.
CC -1- SIMILARITY: IN THE N-TERMINAL SECTION; BELONGS TO THE
CC -1- SODIUM: GALACTOSIDE SYMPORTER FAMILY (SGF).
CC -1- SIMILARITY: Contains 1 Pts EIIA domain.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL: U47655; AAC44113.1; -.
DR HSSP: P45618; 2GPR.
DR InterPro; IPR001927; Na/Gal symport.
DR InterPro; IPR001127; Pts EIIA.
DR Pfam; PF00358; Pts_EIIA_1; 1.
DR ProDom; PD002243; Pts_EIIA_1.
DR TIGRFAMs; TIGR00792; gpb; 1.
DR TIGRFAMs; TIGR00830; PTBA; 1.
DR PROSITE; PS00371; Pts_EIIA_1; 1.
DR PROSITE; PS00872; NA_GALACTOSIDE_SYM; 1.
KM Transmembrane; Sugar transport; Transport; Symport; Phosphorylation;
KW Plasmid.
FT DOMAIN 1 473 SGF DOMAIN.
FT DOMAIN 474 639 EIIA DOMAIN.
FT MOD RES 557 557 PHOSPHORYLATION (BY HPR) (BY SIMILARITY).
FT TRANSMEM 20 40 POTENTIAL.
FT TRANSMEM 59 79 POTENTIAL.
FT TRANSMEM 99 119 POTENTIAL.
FT TRANSMEM 124 144 POTENTIAL.
FT TRANSMEM 176 196 POTENTIAL.
FT TRANSMEM 207 227 POTENTIAL.
FT TRANSMEM 264 284 POTENTIAL.
FT TRANSMEM 294 314 POTENTIAL.
FT TRANSMEM 323 343 POTENTIAL.
FT TRANSMEM 347 367 POTENTIAL.
FT TRANSMEM 398 418 POTENTIAL.
FT TRANSMEM 433 453 POTENTIAL.
SQ SEQUENCE 639 AA; 70153 MW; 25DF28197761B415 CRC64;
Query Match 29.7%; Score 47; DB 1; Length 639;
Best Local Similarity 37.0%; Pred. No. 47;
Matches 10; Conservative 4; Mismatches 13; Indels 0; Gaps 0;
CY 3 TGVLAVGWGEGKGFVPPDDKMSFLFA 29
Db 513 TGIGFADPEEGNLFAPDGVDPFS 539
Search completed: September 30, 2004, 11:10:30
Job time: 12.7714 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: September 30, 2004, 10:58:43 ; Search time 18.5048 Seconds
(without alignments) 150.748 Million cell updates/sec

Title: US-09-988-851A-5
Perfect score: 158
Sequence: 1 VFTGVLAGVWGEGKFPYFPDDKMSFLFA 29

Scoring table: BIOSUM62
Gapop 10.0, Gapext 0.5

Searched: 283366 seqs, 96191526 residues

Total number of hits satisfying chosen parameters: 283366

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :

1: p1r1:*
2: p1r2:*
3: p1r3:*
4: p1r4:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	158	100.0	750	1 WMBEX8	US8 protein - huma
2	56	35.4	226	2 C70383	conserved hypochet
3	53	33.5	3519	2 S43048	polyketide synthas
4	52	32.9	255	2 B64604	conserved hypochet
5	52	32.9	255	2 D71910	hypothetical prote
6	52	32.9	721	2 C64677	probable membrane
7	51	32.3	421	2 C69771	C4-dicarboxylate t
8	50	31.6	357	2 A49024	connexin40 - dog
9	50	31.6	877	2 S58824	probable membrane
10	49	31.0	156	2 B83678	hypothetical prote
11	49	31.0	240	2 T15785	hypothetical prote
12	49	31.0	251	2 T44070	conserved hypochet
13	49	31.0	546	1 STRYEM	histidine-tRNA lig
14	49	31.0	1036	2 T30311	S-layer protein -
15	48.5	30.7	378	2 C64091	adenine glycosylas
16	48.5	30.7	413	2 B81013	tryptophan transpo
17	48.5	30.7	448	2 G83333	probable porin PA2
18	48.5	30.7	506	2 T12819	hypothetical prote
19	48	30.4	277	2 B64548	conserved hypochet
20	48	30.4	277	2 F71960	hypothetical prote
21	48	30.4	293	2 T50940	ditd protein limpo
22	48	30.4	354	2 A71051	conserved hypochet
23	48	30.4	356	2 S56417	hypothetical 40.3k
24	48	30.4	356	2 H91274	hypothetical prote
25	48	30.4	356	2 H86115	hypothetical prote
26	48	30.4	363	2 G95237	conserved hypochet
27	48	30.4	363	2 H88101	conserved hypochet
28	48	30.4	614	2 T66741	hypothetical prote
29	48	30.4	1609	2 A86611	probable outer mem

30	48	30.4	1609	2 H72013	polymorphic membra
31	47.5	30.1	520	2 T24848	histidine-tRNA lig
32	47.5	30.1	525	2 S41763	histidine-tRNA lig
33	47.5	30.1	643	2 T58412	sepa protein - Bac
34	47	29.7	178	2 G69349	ves29-like phospho
35	47	29.7	179	2 D90167	conserved hypochet
36	47	29.7	182	2 B86462	hypothetical prote
37	47	29.7	241	2 S44893	ZK1236.7 protein -
38	47	29.7	651	2 C69374	conserved hypochet
39	47	29.7	856	1 A44963	env polypeptide pr
40	46.5	29.4	208	2 B84920	hypothetical prote
41	46.5	29.4	520	2 B95168	GMP synthase limpo
42	46.5	29.4	520	2 C98034	GMP synthase (glut
43	46	29.1	170	2 G82446	conserved hypochet
44	46	29.1	176	2 AC0635	probable membrane
45	46	29.1	179	2 A64846	probable membrane

ALIGNMENTS

RESULT 1

WMBEX8
US8 protein - human herpesvirus 1 (strain 17)

C/Species: human herpesvirus 1
C/Date: 31-Dec-1989 #sequence_revision 31-Dec-1989 #text_change 16-Jun-2000

C/Accession: C29890; H28133
R/McGeoch, D.J.; Dairymple, M.A.; Dolan, A.; McNab, D.; Perry, L.J.; Taylor, P.; Chalber

J. Virol. 62, 444-453, 1988

A/Title: Structures of herpes simplex virus type 1 genes required for replication of vir

A/Reference number: A93040; M01D:88031053; PMID:2626807

A/Accession: C29890

A/Molecule type: DNA

A/Residues: 1-750 <MCG>

A/Cross-references: GB:M19120; NID:G330226; PIDN:AAA45823.1; PID:G330235

R/McGeoch, D.J.; Dairymple, M.A.; Davison, A.J.; Dolan, A.; Frame, M.C.; McNab, D.; Perry

J. Gen. Virol. 69, 1531-1574, 1988

A/Title: The complete DNA sequence of the long unique region in the genome of herpes sim

A/Reference number: A30083; M01D:88274327; PMID:2839594

A/Accession: H28133

A/Status: nucleic acid sequence not shown; translation not shown

A/Molecule type: DNA

A/Residues: 1-750 <MCG>

A/Cross-references: GB:X14112; NID:G1944536; PIDN:CAA32344.1; PID:G55509; GB:D00317

C/Comment: This protein is required for replication of viral DNA.

C/Genetics:

A/Gene: US8

C/Superfamily: herpesvirus US8 protein

C/Keywords: DNA biosynthesis

Query Match 100.0%; Score 158; DB 1; Length 750;

Best Local Similarity 100.0%; Pred. No. 4e-14; Mismatches 0; Gaps 0;

Matches 29; Conservative 0; Indels 0; Gaps 0;

Qy 1 VFTGVLAGVWGEGKFPYFPDDKMSFLFA 29

Db 722 VFTGVLAGVWGEGKFPYFPDDKMSFLFA 750

RESULT 2

C70383
conserved hypothetical protein aq_963 - Aquifex aeolicus

C/Species: Aquifex aeolicus

C/Date: 08-May-1998 #sequence_revision 08-May-1998 #text_change 05-Nov-1999

C/Accession: C70383

R/Decker, G.; Warren, P.V.; Gaasterland, T.; Young, W.G.; Lenox, A.L.; Graham, D.E.; Ov

V. Nature 392, 353-358, 1998

A/Title: The complete genome of the hyperthermophilic bacterium Aquifex aeolicus.

A/Reference number: A70300; M01D:98196666; PMID:9537320

A/Accession: C70383

A/Status: preliminary; nucleic acid sequence not shown; translation not shown

A/Molecule type: DNA

QY 2 FTGLACVWGE 13
Db 492 FTSLVGLWGTG 503

RESULT 7

C4-dicarboxylate transport protein homolog ydbH - *Bacillus subtilis*
C/Species: *Bacillus subtilis*
C/Date: 05-Dec-1997 #sequence_revision 05-Dec-1997 #text_change 20-Jun-2000
C/Accession: C69771
R/Kunst, F.; Ogasawara, N.; Moszer, I.; Albertini, A.M.; Alloni, G.; Azevedo, V.; Berter
C./Bron, S.; Brouillet, S.; Brusch, C.V.; Caldwell, B.; Capuano, V.; Carter, N.M.; Chd
A./Ehrlich, S.D.; Emerson, P.T.; Entian, K.D.; Errington, J.; Fabret, C.; Ferrari, E.
Nature 390, 249-256, 1997
A/Authors: Foulger, D.; Fritz, C.; Fujita, M.; Fujita, Y.; Fuma, S.; Gallizzi, A.; Gallier
A./Iech, U.; Harwood, C.R.; Henaut, A.; Hilbert, H.; Holsappel, S.; Hosono, S.; Hullo, M.F.
Koetter, P.; Konigstein, G.; Krogh, S.; Kumano, M.; Kurita, K.; Lapidus, A.; Lardinois,
A./Authors: Lauber, J.; Lazarevic, V.; Lee, S.M.; Levine, A.; Liu, H.; Masuda, S.; Maueel
Y., M.; Ogawa, K.; Ogilawa, A.; Oudega, B.; Park, S.H.; Parro, V.; Pohl, T.M.; Portetelli
Riger, M.; Rivolta, C.; Rocha, E.; Roche, B.; Rose, M.; Sadaie, Y.; Sato, T.; Scanlon,
A./Authors: Schleich, S.; Schroeder, R.; Scrofano, F.; Sekiguchi, J.; Sekowska, A.; Serot
Akenchi, M.; Tamakoshi, A.; Tanaka, T.; Terpetra, P.; Tognoni, A.; Tosato, V.; Uchiyama,
T.; Winters, P.; Wipat, A.; Yamamoto, H.; Yamane, K.; Yasumoto, K.; Yata, K.; Yoshida, K.
A/Title: The complete genome sequence of the Gram-positive bacterium *Bacillus subtilis*.
A/Reference number: A69580; MUID:98044033; PMID:9384377
A/Accession: C69771
A/Status: preliminary; nucleic acid sequence not shown; translation not shown
A/Molecule type: DNA
A/Residues: 1-421 <KUN>
A/Cross-references: GB:Z99106; GB:AL009126; NID:G2632653; PIDN:CAB12254.1; PID:G2632747
A/Experimental source: strain 168
C/Genetics:
A/Gene: ydbH
C/Superfamily: C4-dicarboxylate carrier protein

Query Match 32.3%; Score 51; DB 2; Length 421;
Best Local Similarity 43.8%; Pred. No. 21;
Matches 14; Conservative 4; Mismatches 10; Indels 4; Gaps 1;

QY 1 VFTGVLAGV---WGEGRKFPYFPDDKMSFLF 28
Db 151 LFSILFVGVLAAUGERKSYIDFDKVSHV 182

RESULT 8

A49024
connexin40 - dog
C/Species: *Canis lupus familiaris* (dog)
C/Date: 19-Dec-1993 #sequence_revision 18-Nov-1994 #text_change 13-Aug-1999
C/Accession: A49024
R/Kanter, H.L.; Saffitz, J.E.; Beyer, E.C.
Circ. Res. 70, 438-444, 1992
A/Title: Cardiac myocytes express multiple gap junction proteins.
A/Reference number: A49024; MUID:92136511; PMID:1310450
A/Accession: A49024
A/Status: preliminary
A/Molecule type: DNA
A/Residues: 1-357 <KAN>
A/Cross-references: GB:M1347; NID:G163927; PIDN:AAA0838.1; PID:G163928
A/Note: sequence extracted from NCBI backbone (NCBI:81566, NCBI:81570)
C/Superfamily: gap junction protein
C/Keywords: transmembrane protein

Query Match 31.6%; Score 50; DB 2; Length 357;
Best Local Similarity 60.0%; Pred. No. 24;
Matches 9; Conservative 2; Mismatches 4; Indels 0; Gaps 0;

QY 11 GEGGKFPYFPDDKMS 25
Db 273 GPGGKFPYFPFSNKMA 287

RESULT 9
S58824
probable membrane protein YPR194C - yeast (*Saccharomyces cerevisiae*)
N/Alternate names: hypothetical protein P9677.13
C/Species: *Saccharomyces cerevisiae*
C/Date: 28-Nov-1995 #sequence_revision 19-Jan-1996 #text_change 19-Apr-2002
C/Accession: S58824
R/Miller, N.
submitted to the EMBL Data Library, April 1995
A/Description: The sequence of S. cerevisiae cosmid 9677.
A/Reference number: S58816
A/Accession: S58824
A/Molecule type: DNA
A/Residues: 1-877 <ML>
A/Cross-references: EMBL:U25841; NID:G786295; PID:G786296; GSPDB:GN00016; MIPS:YPR194C
C/Genetics:
A/Gene: SGD:OPT2; MIPS:YPR194C
A/Cross-references: SGD:S0006398
A/Map position: 16R
C/Superfamily: *Saccharomyces cerevisiae* probable membrane protein YUL212C

C/Keywords: transmembrane protein
F/278-294/Domain: transmembrane #status predicted <TW1>
F/409-425/Domain: transmembrane #status predicted <TW2>
F/485-501/Domain: transmembrane #status predicted <TW3>
F/558-574/Domain: transmembrane #status predicted <TW4>
F/586-602/Domain: transmembrane #status predicted <TW5>
F/669-685/Domain: transmembrane #status predicted <TW6>
F/734-750/Domain: transmembrane #status predicted <TW7>
F/816-832/Domain: transmembrane #status predicted <TW8>

Query Match 31.6%; Score 50; DB 2; Length 877;
Best Local Similarity 41.7%; Pred. No. 62;
Matches 10; Conservative 3; Mismatches 11; Indels 0; Gaps 0;

QY 4 GTLACVWGEGRKFPYFPDDKMSFL 27
Db 746 GIFFGVKRWKGFYFPDDMLFV 769

RESULT 10

E83678
hypothetical protein BH0229 [imported] - *Bacillus halodurans* (strain C-125)
C/Species: *Bacillus halodurans*
C/Date: 01-Dec-2000 #sequence_revision 01-Dec-2000 #text_change 15-Jun-2001
C/Accession: E83678
R/Takami, H.; Nakasone, K.; Takaki, Y.; Maeno, G.; Sasaki, R.; Masui, N.; Fujii, F.; Hiran
Nucleic Acids Res. 28, 4317-4331, 2000
A/Title: Complete genome sequence of the alkaliphilic bacterium *Bacillus halodurans* and
A/Reference number: A83650; MUID:20512582; PMID:11058132
A/Accession: E83678
A/Status: preliminary
A/Molecule type: DNA
A/Residues: 1-156 <STO>
A/Cross-references: GB:AP001507; GB:BA000004; NID:G10172612; PIDN:BA03948.1; GSPDB:GN001
A/Experimental source: strain C-125
C/Genetics:
C/Superfamily: *Bacillus subtilis* hypothetical protein ywID

Query Match 31.0%; Score 49; DB 2; Length 156;
Best Local Similarity 64.3%; Pred. No. 14;
Matches 9; Conservative 2; Mismatches 1; Indels 2; Gaps 1;

QY 10 WGEGRKFPYFPDDKMS 23
Db 77 WGEGR--VFPDDKMS 88

RESULT 11
T15785
hypothetical protein C39D10.6 - *Caenorhabditis elegans*

Fri Oct 1 12:26:20 2004

us-09-988-851a-5.rpr

Page 5

A; Status: nucleic acid sequence not shown; translation not shown
A; Molecule type: DNA

A;Molecule type: DNA

A;Residues: 1-378 <TIGR>

A; Cross-references: GB:U

C;Function:

A/Description: catalyzes the excision of adenine from a guanine-adenine mispair;
C:Superfamily: A/C-specific adenine glycosylase

C/Superfamily: A/G-specific adenine glycosylase

C;Keywords: DNA repair

Query Match	30.7%;	Score 48.5;	DB 2;	Length 378;
-------------	--------	-------------	-------	-------------

Best Local Similarity 50.0%; Pred. No. 42;

Matches 11; Conservative 4; Mismatches 4; Indels 3; Gaps 2;

QY 7 AGVMGEGGKFVYP-FDDKMSFL 27

Db 257 SGLW--GGLFCFPQFEDKSSLL 276

```
Search completed: September 30, 2004, 11:09:45
Job time : 20.5048 secs
```

Job time : 20.5048 secs

This Page Blank (uspto)

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: September 30, 2004, 10:58:43 ; Search time 75.1238 Seconds
(without alignments)
109.072 Million cell updates/sec

Title: US-09-988-851A-5
Perfect score: 158
Sequence: 1 VFTGLAGVWGEGKFPYPDDKMSFLPA 29

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 1586107 seqs, 282547505 residues

Total number of hits satisfying chosen parameters: 1586107

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : A_Geneseq_29JAr04:*
1: Geneseqp1980s:*
2: Geneseqp1990s:*
3: Geneseqp2000s:*
4: Geneseqp2001s:*
5: Geneseqp2002s:*
6: Geneseqp2003as:*
7: Geneseqp2003bs:*
8: Geneseqp2004s:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	158	100.0	29	2	AAW53265 Herpes si
2	148	99.7	27	2	AAW53272 Herpes si
3	137	86.7	25	2	AAW53271 Herpes si
4	131	82.9	386	2	AAW72064 HSV-2 str
5	131	82.9	408	2	AAW72138 HSV-2 str
6	131	82.9	750	2	AAW72176 HSV-2 str
7	129	81.6	23	2	AAW53270 Herpes si
8	92	58.2	20	2	AAW53266 Herpes si
9	82	51.9	15	2	AAW53273 Herpes si
10	66	41.8	12	2	AAW53269 Herpes si
11	61	38.6	60	4	AAW84164 Human imm
12	54.5	34.5	373	7	ADC97033 E. faeciu
13	53	33.5	134	4	ABG20173 Novel hum
14	53	33.5	1580	2	AAW22609 Platenoli
15	53	33.5	1580	2	AAW23719 Platenoli
16	53	33.5	1891	2	AAW22610 Platenoli
17	53	33.5	1891	2	AAW23720 Platenoli
18	53	33.5	3519	3	AAW92709 S. antibi
19	53	33.5	4150	3	AAW92707 S. antibi
20	52	32.9	255	2	AAW71478 Helicobac
21	52	32.9	255	4	AAW46309 H. pylori
22	52	32.9	255	4	AAW46359 H. pylori
23	51	32.3	4545	2	AAW22611 Hydril sr
24	51	32.3	4550	2	AAW22606 Platenoli
25	51	32.3	4550	2	AAW23716 Platenoli

26	51	32.3	7068	4	AAE10142 Streptomy
27	51	32.3	9477	4	AAE10144 Streptomy
28	50.5	32.0	99	6	ABP81046 N. gonorr
29	50	31.6	85	4	ABP12216 Human sec
30	50	31.6	117	7	AAU17365 Novel sig
31	50	31.6	117	7	ADP94073 Human nov
32	50	31.6	129	3	AAQ01459 Human sec
33	50	31.6	204	4	AAE68512 Human GTP
34	50	31.6	204	5	ABP69807 Human pol
35	50	31.6	262	6	ADA35285 Acinetoba
36	50	31.6	389	3	AAAB03761 BsrI met
37	50	31.6	877	3	AAW96168 Saccharom
38	50	31.6	3722	2	AAE10145 Cephalosp
39	50	31.6	3816	3	AAW92708 S. antibi
40	49	31.0	251	6	ABW73089 Staphyloc
41	49	31.0	480	4	AAW92386 C. glutam
42	49	31.0	497	4	AAW79227 Corynebac
43	49	31.0	526	2	AAW25222 Yeast his
44	49	31.0	528	7	ADD13623 C. glutam
45	49	31.0	1784	6	AAE37003 Micromono

ALIGNMENTS

RESULT 1	AAW53265	AAW53265 standard; peptide; 29 AA.
ID	AAW53265	
XX	AAW53265;	
AC		
DT	01-JUL-1998 (first entry)	
XX		
DE	Herpes simplex virus type 1 antiviral agent peptide 5.	
XX		
KM	Herpes simplex virus type 1; HSV-1; antiviral; inhibition; infection; replication; Uls; POL.	
XX		
OS	Synthetic.	
OS	Herpes simplex virus unknown type.	
PN	W09804707-A1.	
XX		
PD	05-FEB-1998.	
XX		
PF	28-JUL-1997; 97WO-GB002025.	
XX		
PR	26-JUL-1996; 96GB-00015730.	
XX		
PA	(MEDT-) MEDICAL RES COUNCIL.	
XX		
PI	Marsden HS, Stow ND, Mclean GW;	
XX		
DR	WPI; 1998-130695/12.	
XX		
PT	Antiviral agent capable of inhibiting herpes virus replication - disrupts association between herpes simplex virus type 1 Uls and POL.	
XX		
PS	Claim 5; Page 57; 83pp; English.	
CC	The present sequence represents an antiviral agent peptide capable of disrupting herpes virus replication. The antiviral agent is capable of disrupting the association between Uls and POL (UL30), where Uls and POL are respectively defined as Uls and POL of herpes simplex virus type 1 (HSV-1) together with homologues in other herpes viruses. The present invention also describes an assay to determine the ability of a test substance to interfere with the association of Uls and POL. The assay comprises: (a) exposing a 1st viral component to a test substance followed by a 2nd viral component, or exposing a 1st viral component to a 2nd viral component followed by a test substance; (b) washing to remove any 2nd viral component and/or test substance not associated with the 1st viral component; and (c) detecting the presence, and optionally determining the amount, of 2nd viral component associated with the 1st	

CC viral component. The antiviral agent can be used to combat herpes virus
 CC replication or infection
 XX
 SQ Sequence 29 AA;

Query Match 100.0%; Score 158; DB 2; Length 29;
 Best Local Similarity 100.0%; Pred. No. 7, 8e-15;
 Matches 29; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

DB 1 VFTGVLAGWGEKGKFPVFPDDKMSFLFA 29
 1 VFTGVLAGWGEKGKFPVFPDDKMSFLFA 29

RESULT 2

AAW53272
 ID AAW53272 standard; peptide; 27 AA.

AC AAW53272;

DT 01-JUL-1998 (first entry)

DE Herpes simplex virus type 1 antiviral peptide 4.

KW Herpes simplex virus type 1; HSV-1; antiviral; inhibition; infection;
 KM replication; UL8; POL.

OS Synthetic.
 OS Herpes simplex virus unknown type.

PN MO9804707-A1.

PD 05-FEB-1998.

PF 28-JUL-1997; 97WO-GB002025.

PR 26-JUL-1996; 96GB-00015730.

PS (MEDI-) MEDICAL RES COUNCIL.

PI Marsden HS, Stow ND, McLean GW;

DR WPI; 1998-130695/12.

PT Antiviral agent capable of inhibiting herpes virus replication - disrupts
 association between herpes simplex virus type 1 UL8 and POL.

PS Example 1; Page 43; 83pp; English.

CC The present sequence represents a peptide used in an example of the
 CC present invention. The present invention describes an antiviral agent
 CC capable of combating herpes virus replication. The antiviral agent is
 CC capable of disrupting the association between UL8 and POL (UL30), where
 CC UL8 and POL are respectively defined as UL8 and POL of herpes simplex
 CC virus type 1 (HSV-1) together with homologues in other herpes viruses.
 CC The present invention also describes an assay to determine the ability of
 CC a test substance to interfere with the association of UL8 and POL. The
 CC assay comprises: (a) exposing a 1st viral component to a test substance
 CC followed by a 2nd viral component, or exposing a 1st viral component to a
 CC 2nd viral component followed by a test substance; (b) washing to remove
 CC any 2nd viral component and/or test substance not associated with the 1st
 CC viral component; and (c) detecting the presence, and optionally
 CC determining the amount, of 2nd viral component associated with the 1st
 CC viral component. The antiviral agent can be used to combat herpes virus
 CC replication or infection

SQ Sequence 27 AA;

Query Match 93.7%; Score 148; DB 2; Length 27;
 Best Local Similarity 100.0%; Pred. No. 1, 8e-13;
 Matches 27; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3 TGVLAGWGEKGKFPVFPDDKMSFLFA 29

DB 1 TGVLAGWGEKGKFPVFPDDKMSFLFA 27

RESULT 3

AAW53271
 ID AAW53271 standard; peptide; 25 AA.

AC AAW53271;

DT 01-JUL-1998 (first entry)

DE Herpes simplex virus type 1 antiviral peptide 3.

KW Herpes simplex virus type 1; HSV-1; antiviral; inhibition; infection;
 KM replication; UL8; POL.

OS Synthetic.
 OS Herpes simplex virus unknown type.

PN MO9804707-A1.

PD 05-FEB-1998.

PF 28-JUL-1997; 97WO-GB002025.

PR 26-JUL-1996; 96GB-00015730.

PS (MEDI-) MEDICAL RES COUNCIL.

PI Marsden HS, Stow ND, McLean GW;

DR WPI; 1998-130695/12.

PT Antiviral agent capable of inhibiting herpes virus replication - disrupts
 association between herpes simplex virus type 1 UL8 and POL.

PS Example 1; Page 43; 83pp; English.

CC The present sequence represents a peptide used in an example of the
 CC present invention. The present invention describes an antiviral agent
 CC capable of combating herpes virus replication. The antiviral agent is
 CC capable of disrupting the association between UL8 and POL (UL30), where
 CC UL8 and POL are respectively defined as UL8 and POL of herpes simplex
 CC virus type 1 (HSV-1) together with homologues in other herpes viruses.
 CC The present invention also describes an assay to determine the ability of
 CC a test substance to interfere with the association of UL8 and POL. The
 CC assay comprises: (a) exposing a 1st viral component to a test substance
 CC followed by a 2nd viral component, or exposing a 1st viral component to a
 CC 2nd viral component followed by a test substance; (b) washing to remove
 CC any 2nd viral component and/or test substance not associated with the 1st
 CC viral component; and (c) detecting the presence, and optionally
 CC determining the amount, of 2nd viral component associated with the 1st
 CC viral component. The antiviral agent can be used to combat herpes virus
 CC replication or infection

SQ Sequence 25 AA;

Query Match 86.7%; Score 137; DB 2; Length 25;
 Best Local Similarity 100.0%; Pred. No. 5, 7e-12;
 Matches 25; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 5 VLAGWGEKGKFPVFPDDKMSFLFA 29
 1 VLAGWGEKGKFPVFPDDKMSFLFA 25

RESULT 4

AAW72064
 ID AAW72064 standard; protein; 386 AA.

AC AAW72064;

QY 3 TGVLAGWGEKGKFPVFPDDKMSFLFA 29

```

DT 18-DEC-1998 (first entry)
XX
DE HSV-2 strain SBS Contig ID 93 ORF#2 protein.
XX
KM HSV-2 strain SBS; immunological response induction; therapy;
KM antiviral identification; viral protein inhibitor.
XX
OS Herpes simplex virus 2.
XX
PN WO9820016-A1.
XX
PD 14-MAY-1998.
XX
PF 31-OCT-1997; 97WO-US020016.
XX
PR 04-NOV-1996; 96US-0030279P.
PR 09-JUN-1997; 97US-0049018P.
XX
PA (SMIK ) SMITHKLINE BEECHAM CORP.
XX
PI Esser KM, Chan JY, Dabrowski-Amara CE, Delvecchio AM, Dillon SB;
PI Leary JT;
XX
DR WPI; 1998-286847/25.
DR N-PSDB; AAV62147.
XX
PT Herpes simplex virus type-2 sequences - useful in, e.g. prevention and
PT treatment of infection or inducing immunological response in mammal.
XX
PS Claim 10; Page 65; 748pp; English.
XX
CC This sequence represents a Herpes simplex virus type-2 (HSV-2) protein
CC sequence of the invention. This sequence was isolated from a HSV-2 strain
CC SB5 (deposited as ATCC VR-2546) DNA fragment designated Contig ID 93.
CC Based on homology, this sequence is a US protein. The proteins can be
CC used for the treatment or prevention of disease, to induce an
CC immunological response in a mammal or to identify inhibitors, activators
CC or novel antivirals. Antagonists of the proteins can be used to inhibit a
CC viral polypeptide. The DNA sequence or a vector containing it can also be
CC used to induce an immunological response in a mammal
XX
SQ Sequence 386 AA;

Query Match 82.9%; Score 131; DB 2; Length 386;
Best Local Similarity 82.1%; Pred. No. 6.4e-10;
Matches 23; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

QY 1 VFTGVLGVGEGGKFFVFPDDKMSFLF 28
   ||||| ||||| ||||| ||||| |||||
Db 358 VFTGVLGVGEGGKFFVFPDDKMSFLF 385

RESULT 5
AAW72138
ID AAW72138 standard; protein; 408 AA.
XX
AC AAW72138;
XX
DT 23-DEC-1998 (first entry)
XX
DE HSV-2 strain SBS Contig ID 18 ORF#3 protein.
XX
KM HSV-2 strain SBS; immunological response induction; therapy;
KM antiviral identification; viral protein inhibitor.
XX
OS Herpes simplex virus 2.
XX
PN WO9820016-A1.
XX
PD 14-MAY-1998.
XX
PF 31-OCT-1997; 97WO-US020016.
XX
PR 04-NOV-1996; 96US-0030279P.
PR 09-JUN-1997; 97US-0049018P.
XX
PA (SMIK ) SMITHKLINE BEECHAM CORP.
XX
PI Esser KM, Chan JY, Dabrowski-Amara CE, Delvecchio AM, Dillon SB;
PI Leary JT;
XX
DR WPI; 1998-286847/25.
DR N-PSDB; AAV62147.
XX
PT Herpes simplex virus type-2 sequences - useful in, e.g. prevention and
PT treatment of infection or inducing immunological response in mammal.
XX
PS Claim 10; Page 97; 748pp; English.
XX
CC This sequence represents a Herpes simplex virus type-2 (HSV-2) protein
CC sequence of the invention. This sequence was isolated from a HSV-2 strain
CC SB5 (deposited as ATCC VR-2546) DNA fragment designated Contig ID 18. The
CC proteins can be used for the treatment or prevention of disease, to
CC induce an immunological response in a mammal or to identify inhibitors,
CC activators or novel antivirals. Antagonists of the proteins can be used
CC to inhibit a viral polypeptide. The DNA sequence or a vector containing
CC it can also be used to induce an immunological response in a mammal
XX
SQ Sequence 408 AA;

Query Match 82.9%; Score 131; DB 2; Length 408;
Best Local Similarity 82.1%; Pred. No. 6.8e-10;
Matches 23; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

QY 1 VFTGVLGVGEGGKFFVFPDDKMSFLF 28
   ||||| ||||| ||||| ||||| |||||
Db 380 VFTGVLGVGEGGKFFVFPDDKMSFLF 407

RESULT 6
AAW72176
ID AAW72176 standard; protein; 750 AA.
XX
AC AAW72176;
XX
DT 13-JUN-1999 (first entry)
XX
DE HSV-2 strain SBS Contig ID 15 ORF#8 protein.
XX
KM HSV-2 strain SBS; immunological response induction; therapy;
KM antiviral identification; viral protein inhibitor.
XX
OS Herpes simplex virus 2.
XX
PN WO9820016-A1.
XX
PD 14-MAY-1998.
XX
PF 31-OCT-1997; 97WO-US020016.
XX
PR 04-NOV-1996; 96US-0030279P.
PR 09-JUN-1997; 97US-0049018P.
XX
PA (SMIK ) SMITHKLINE BEECHAM CORP.
XX
PI Esser KM, Chan JY, Dabrowski-Amara CE, Delvecchio AM, Dillon SB;
PI Leary JT;
XX

```

```

PD 14-MAY-1998.
XX
PF 31-OCT-1997; 97WO-US020016.
XX
PR 04-NOV-1996; 96US-0030279P.
PR 09-JUN-1997; 97US-0049018P.
XX
PA (SMIK ) SMITHKLINE BEECHAM CORP.
XX
PI Esser KM, Chan JY, Dabrowski-Amara CE, Delvecchio AM, Dillon SB;
PI Leary JT;
XX
DR WPI; 1998-286847/25.
DR N-PSDB; AAV62147.
XX
PT Herpes simplex virus type-2 sequences - useful in, e.g. prevention and
PT treatment of infection or inducing immunological response in mammal.
XX
PS Claim 10; Page 97; 748pp; English.
XX
CC This sequence represents a Herpes simplex virus type-2 (HSV-2) protein
CC sequence of the invention. This sequence was isolated from a HSV-2 strain
CC SB5 (deposited as ATCC VR-2546) DNA fragment designated Contig ID 18. The
CC proteins can be used for the treatment or prevention of disease, to
CC induce an immunological response in a mammal or to identify inhibitors,
CC activators or novel antivirals. Antagonists of the proteins can be used
CC to inhibit a viral polypeptide. The DNA sequence or a vector containing
CC it can also be used to induce an immunological response in a mammal
XX
SQ Sequence 408 AA;

Query Match 82.9%; Score 131; DB 2; Length 408;
Best Local Similarity 82.1%; Pred. No. 6.8e-10;
Matches 23; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

QY 1 VFTGVLGVGEGGKFFVFPDDKMSFLF 28
   ||||| ||||| ||||| ||||| |||||
Db 380 VFTGVLGVGEGGKFFVFPDDKMSFLF 407

RESULT 6
AAW72176
ID AAW72176 standard; protein; 750 AA.
XX
AC AAW72176;
XX
DT 13-JUN-1999 (first entry)
XX
DE HSV-2 strain SBS Contig ID 15 ORF#8 protein.
XX
KM HSV-2 strain SBS; immunological response induction; therapy;
KM antiviral identification; viral protein inhibitor.
XX
OS Herpes simplex virus 2.
XX
PN WO9820016-A1.
XX
PD 14-MAY-1998.
XX
PF 31-OCT-1997; 97WO-US020016.
XX
PR 04-NOV-1996; 96US-0030279P.
PR 09-JUN-1997; 97US-0049018P.
XX
PA (SMIK ) SMITHKLINE BEECHAM CORP.
XX
PI Esser KM, Chan JY, Dabrowski-Amara CE, Delvecchio AM, Dillon SB;
PI Leary JT;
XX

```

DR WPI; 1998-286847/25.
DR N-PSDB; AAV62176.

PT Herpes simplex virus type-2 sequences - useful in, e.g. prevention and
XX treatment of infection or inducing immunological response in mammal.

PS Claim 10; Page 112; 748bp; English.

XX This sequence represents a Herpes simplex virus type-2 (HSV-2) protein
CC sequence of the invention. This sequence was isolated from a HSV-2 strain
CC SB5 (deposited as ATCC VR-2546) DNA fragment designated Contig ID 15. The
CC proteins can be used for the treatment or prevention of disease, to
CC induce an immunological response in a mammal or to identify inhibitors,
CC activators or novel antivirals. Antagonists of the proteins can be used
CC to inhibit a viral polypeptide. The DNA sequence or a vector containing
CC it can also be used to induce an immunological response in a mammal
XX

SQ Sequence 750 AA;

Query Match 82.9%; Score 131; DB 2; Length 750;
Best Local Similarity 82.1%; Pred. No. 1.3e-05;
Matches 23; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

QY 1 VFTGVLGVWGGGKFFYPDPDKMSFLP 28
DB 722 VFTGVLGVWGGGKFFYPDPDKMSFLP 749

RESULT 7
AAW53270
ID AAW53270 standard; peptide; 23 AA.

XX AAW53270;

XX 01-JUL-1998 (first entry)

XX Herpes simplex virus type 1 antiviral peptide 2.

XX Herpes simplex virus type 1; HSV-1; antiviral; inhibition; infection;
KW replication; UL8; POL.

XX Synthetic

OS Herpes simplex virus unknown type.

XX WO9804707-A1.

XX 05-FEB-1998.

XX 28-JUL-1997; 97WO-GB002025.

XX 26-JUL-1996; 96GB-00015730.

XX (MEDI-) MEDICAL RES COUNCIL.

XX Marsden HS, Stow ND, McLean GW;

XX WPI; 1998-130695/12.

XX Antiviral agent capable of inhibiting herpes virus replication - disrupts
PT association between herpes simplex virus type 1 UL8 and POL.

XX Example 1; Page 43; 83pp; English.

XX The present sequence represents a peptide used in an example of the
CC present invention. The present invention describes an antiviral agent
CC capable of combating herpes virus replication. The antiviral agent is
CC capable of disrupting the association between UL8 and POL (UL30), where
CC UL8 and POL are respectively defined as UL8 and POL of herpes simplex
CC virus type 1 (HSV-1) together with homologues in other herpes viruses.
CC The present invention also describes an assay to determine the ability of
CC a test substance to interfere with the association of UL8 and POL. The
CC assay comprises: (a) exposing a 1st viral component to a test substance
CC followed by a 2nd viral component, or exposing a 1st viral component to a

CC 2nd viral component followed by a test substance; (b) washing to remove
CC any 2nd viral component and/or test substance not associated with the 1st
CC viral component; and (c) detecting the presence, and optionally
CC determining the amount, of 2nd viral component associated with the 1st
CC viral component. The antiviral agent can be used to combat herpes virus
CC replication or infection
XX

SQ Sequence 23 AA;

Query Match 81.6%; Score 129; DB 2; Length 23;
Best Local Similarity 100.0%; Pred. No. 6.8e-11;
Matches 23; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 7 AGWGGGKFFYPDPDKMSFLPFA 29
DB 1 AGWGGGKFFYPDPDKMSFLPFA 23

RESULT 8
AAW53266
ID AAW53266 standard; peptide; 20 AA.

XX AAW53266;

XX 01-JUL-1998 (first entry)

XX Herpes simplex virus type 1 antiviral agent peptide 7.

XX Herpes simplex virus type 1; HSV-1; antiviral; inhibition; infection;
KW replication; UL8; POL.

XX Synthetic

OS Herpes simplex virus unknown type.

XX WO9804707-A1.

XX 05-FEB-1998.

XX 28-JUL-1997; 97WO-GB002025.

XX 26-JUL-1996; 96GB-00015730.

XX (MEDI-) MEDICAL RES COUNCIL.

XX Marsden HS, Stow ND, McLean GW;

XX WPI; 1998-130695/12.

XX Antiviral agent capable of inhibiting herpes virus replication - disrupts
PT association between herpes simplex virus type 1 UL8 and POL.

XX Claim 5; Page 57; 83pp; English.

XX The present sequence represents an antiviral agent peptide capable of
CC combating herpes virus replication. The antiviral agent is capable of
CC disrupting the association between UL8 and POL (UL30), where UL8 and POL
CC are respectively defined as UL8 and POL of herpes simplex virus type 1
CC (HSV-1) together with homologues in other herpes viruses. The present
CC invention also describes an assay to determine the ability of a test
CC substance to interfere with the association of UL8 and POL. The assay
CC comprises: (a) exposing a 1st viral component to a test substance
CC followed by a 2nd viral component, or exposing a 1st viral component to a
CC 2nd viral component followed by a test substance; (b) washing to remove
CC any 2nd viral component and/or test substance not associated with the 1st
CC viral component; and (c) detecting the presence, and optionally
CC determining the amount, of 2nd viral component associated with the 1st
CC viral component. The antiviral agent can be used to combat herpes virus
CC replication or infection
XX

SQ Sequence 20 AA;

Query Match 58.2%; Score 92; DB 2; Length 20;
Best Local Similarity 100.0%; Pred. No. 8.4e-06;

Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 VFTGVLGVWGEGRKFV 17
DB 4 VFTGVLGVWGEGRKFV 20

RESULT 9

AAW53273 ID AAW53273 standard; peptide, 15 AA.

XX AAW53273;

XX 01-JUL-1998 (first entry)

XX Herpes simplex virus type 1 antiviral peptide 6.

XX Herpes simplex virus type 1; HSV-1; antiviral; inhibition; infection;

XX replication; UL8; POL.

XX Synthetic.

XX Herpes simplex virus unknown type.

XX WO9804707-A1.

XX 05-FEB-1998.

XX 28-JUL-1997; 97WO-GB002025.

XX 26-JUL-1996; 96GB-00015730.

XX (MEDI-) MEDICAL RES COUNCIL.

XX Marsden HS, Slow ND, Mclean GW;

XX WPI; 1998-130695/12.

XX Antiviral agent capable of inhibiting herpes virus replication - disrupts association between herpes simplex virus type 1 UL8 and POL.

XX Example 1; Page 43; 83pp; English.

XX The present sequence represents a peptide used in an example of the present invention. The present invention describes an antiviral agent capable of combatting herpes virus replication. The antiviral agent is capable of disrupting the association between UL8 and POL (UL30), where UL8 and POL are respectively defined as UL8 and POL of herpes simplex virus type 1 (HSV-1) together with homologues in other herpes viruses. The present invention also describes an assay to determine the ability of a test substance to interfere with the association of UL8 and POL. The assay comprises: (a) exposing a 1st viral component to a test substance followed by a 2nd viral component, or exposing a 1st viral component to a 2nd viral component followed by a test substance; (b) washing to remove any 2nd viral component and/or test substance not associated with the 1st viral component; and (c) detecting the presence, and optionally determining the amount, of 2nd viral component associated with the 1st viral component. The antiviral agent can be used to combat herpes virus replication or infection

XX Sequence 15 AA;

Query Match 51.9%; Score 82; DB 2; Length 15;

Best Local Similarity 100.0%; Pred. No. 0.00016;

Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3 TGVLAGVWGEGRKFV 17

DB 1 TGVLAGVWGEGRKFV 15

RESULT 10

AAW53269 ID AAW53269 standard; peptide, 12 AA.

XX AAW53269;

XX 01-JUL-1998 (first entry)

XX Herpes simplex virus type 1 antiviral peptide 1.

XX Herpes simplex virus type 1; HSV-1; antiviral; inhibition; infection;

XX replication; UL8; POL.

XX Synthetic.

XX Herpes simplex virus unknown type.

XX WO9804707-A1.

XX 05-FEB-1998.

XX 28-JUL-1997; 97WO-GB002025.

XX 26-JUL-1996; 96GB-00015730.

XX (MEDI-) MEDICAL RES COUNCIL.

XX Marsden HS, Slow ND, Mclean GW;

XX WPI; 1998-130695/12.

XX Antiviral agent capable of inhibiting herpes virus replication - disrupts association between herpes simplex virus type 1 UL8 and POL.

XX Example 1; Page 43; 83pp; English.

XX The present sequence represents a peptide used in an example of the present invention. The present invention describes an antiviral agent capable of combatting herpes virus replication. The antiviral agent is capable of disrupting the association between UL8 and POL (UL30), where UL8 and POL are respectively defined as UL8 and POL of herpes simplex virus type 1 (HSV-1) together with homologues in other herpes viruses. The present invention also describes an assay to determine the ability of a test substance to interfere with the association of UL8 and POL. The assay comprises: (a) exposing a 1st viral component to a test substance followed by a 2nd viral component, or exposing a 1st viral component to a 2nd viral component followed by a test substance; (b) washing to remove any 2nd viral component and/or test substance not associated with the 1st viral component; and (c) detecting the presence, and optionally determining the amount, of 2nd viral component associated with the 1st viral component. The antiviral agent can be used to combat herpes virus replication or infection

XX Sequence 12 AA;

Query Match 41.8%; Score 66; DB 2; Length 12;

Best Local Similarity 100.0%; Pred. No. 0.021;

Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 18 YPDDXMSFLFA 29

DB 1 YPDDXMSFLFA 12

RESULT 11

AAW84164 ID AAW84164 standard; protein, 60 AA.

XX AAW84164;

XX 07-NOV-2001 (first entry)

XX Human immune/haematopoietic antigen SEQ ID NO:11757.

XX Human; immune; haematopoietic; immune/haematopoietic antigen; cancer;

XX cytostatic; gene therapy; vaccine; metastasis.

OS Homo sapiens.
 XX
 EN WO200157182-A2.
 XX
 PD 09-AUG-2001.
 XX
 PF 17-JAN-2001; 2001WO-US001354.
 XX
 PR 31-JAN-2000; 2000US-0179065P.
 PR 04-FEB-2000; 2000US-0180628P.
 PR 24-FEB-2000; 2000US-0184664P.
 PR 02-MAR-2000; 2000US-0186350P.
 PR 16-MAR-2000; 2000US-0189874P.
 PR 17-MAR-2000; 2000US-0190076P.
 PR 18-APR-2000; 2000US-0198123P.
 PR 19-MAY-2000; 2000US-020515P.
 PR 07-JUN-2000; 2000US-0209467P.
 PR 28-JUN-2000; 2000US-0214886P.
 PR 30-JUN-2000; 2000US-0215135P.
 PR 07-JUL-2000; 2000US-0216647P.
 PR 07-JUL-2000; 2000US-0216880P.
 PR 11-JUL-2000; 2000US-0217487P.
 PR 11-JUL-2000; 2000US-0217496P.
 PR 14-JUL-2000; 2000US-0218290P.
 PR 26-JUL-2000; 2000US-0220963P.
 PR 26-JUL-2000; 2000US-0220964P.
 PR 14-AUG-2000; 2000US-0224518P.
 PR 14-AUG-2000; 2000US-0224519P.
 PR 14-AUG-2000; 2000US-0225213P.
 PR 14-AUG-2000; 2000US-0225214P.
 PR 14-AUG-2000; 2000US-0225266P.
 PR 14-AUG-2000; 2000US-0225267P.
 PR 14-AUG-2000; 2000US-0225268P.
 PR 14-AUG-2000; 2000US-0225270P.
 PR 14-AUG-2000; 2000US-0225447P.
 PR 14-AUG-2000; 2000US-0225757P.
 PR 14-AUG-2000; 2000US-0225758P.
 PR 14-AUG-2000; 2000US-0225759P.
 PR 18-AUG-2000; 2000US-0226279P.
 PR 22-AUG-2000; 2000US-0226681P.
 PR 22-AUG-2000; 2000US-0226868P.
 PR 22-AUG-2000; 2000US-0227182P.
 PR 23-AUG-2000; 2000US-0227009P.
 PR 30-AUG-2000; 2000US-0228924P.
 PR 01-SEP-2000; 2000US-0229287P.
 PR 01-SEP-2000; 2000US-0229343P.
 PR 01-SEP-2000; 2000US-0229344P.
 PR 01-SEP-2000; 2000US-0229345P.
 PR 05-SEP-2000; 2000US-0229509P.
 PR 05-SEP-2000; 2000US-0229513P.
 PR 06-SEP-2000; 2000US-0230437P.
 PR 06-SEP-2000; 2000US-0230438P.
 PR 08-SEP-2000; 2000US-0231242P.
 PR 08-SEP-2000; 2000US-0231243P.
 PR 08-SEP-2000; 2000US-0231244P.
 PR 08-SEP-2000; 2000US-0231413P.
 PR 08-SEP-2000; 2000US-0231414P.
 PR 08-SEP-2000; 2000US-0232080P.
 PR 08-SEP-2000; 2000US-0232081P.
 PR 12-SEP-2000; 2000US-0231968P.
 PR 14-SEP-2000; 2000US-0232397P.
 PR 14-SEP-2000; 2000US-0232398P.
 PR 14-SEP-2000; 2000US-0232399P.
 PR 14-SEP-2000; 2000US-0232400P.
 PR 14-SEP-2000; 2000US-0233063P.
 PR 14-SEP-2000; 2000US-0233064P.
 PR 14-SEP-2000; 2000US-0233065P.
 PR 21-SEP-2000; 2000US-0234222P.
 PR 21-SEP-2000; 2000US-0234274P.
 PR 25-SEP-2000; 2000US-0234997P.
 PR 25-SEP-2000; 2000US-0234998P.
 PR 26-SEP-2000; 2000US-0235484P.

PR 27-SEP-2000; 2000US-0235834P.
 PR 27-SEP-2000; 2000US-0235836P.
 PR 29-SEP-2000; 2000US-0236327P.
 PR 29-SEP-2000; 2000US-0236367P.
 PR 29-SEP-2000; 2000US-0236368P.
 PR 29-SEP-2000; 2000US-0236369P.
 PR 29-SEP-2000; 2000US-0236370P.
 PR 02-OCT-2000; 2000US-0237037P.
 PR 02-OCT-2000; 2000US-0237038P.
 PR 02-OCT-2000; 2000US-0237039P.
 PR 02-OCT-2000; 2000US-0237040P.
 PR 13-OCT-2000; 2000US-0239933P.
 PR 13-OCT-2000; 2000US-0239937P.
 PR 20-OCT-2000; 2000US-0240960P.
 PR 20-OCT-2000; 2000US-0241221P.
 PR 20-OCT-2000; 2000US-0241785P.
 PR 20-OCT-2000; 2000US-0241786P.
 PR 20-OCT-2000; 2000US-0241787P.
 PR 20-OCT-2000; 2000US-0241808P.
 PR 20-OCT-2000; 2000US-0241809P.
 PR 20-OCT-2000; 2000US-0241826P.
 PR 20-OCT-2000; 2000US-0241828P.
 PR 01-NOV-2000; 2000US-0244617P.
 PR 08-NOV-2000; 2000US-0246474P.
 PR 08-NOV-2000; 2000US-0246475P.
 PR 08-NOV-2000; 2000US-0246476P.
 PR 08-NOV-2000; 2000US-0246477P.
 PR 08-NOV-2000; 2000US-0246478P.
 PR 08-NOV-2000; 2000US-0246523P.
 PR 08-NOV-2000; 2000US-0246524P.
 PR 08-NOV-2000; 2000US-0246525P.
 PR 08-NOV-2000; 2000US-0246526P.
 PR 08-NOV-2000; 2000US-0246527P.
 PR 08-NOV-2000; 2000US-0246528P.
 PR 08-NOV-2000; 2000US-0246532P.
 PR 08-NOV-2000; 2000US-0246608P.
 PR 08-NOV-2000; 2000US-0246610P.
 PR 08-NOV-2000; 2000US-0246611P.
 PR 08-NOV-2000; 2000US-0246613P.
 PR 17-NOV-2000; 2000US-0249207P.
 PR 17-NOV-2000; 2000US-0249208P.
 PR 17-NOV-2000; 2000US-0249209P.
 PR 17-NOV-2000; 2000US-0249210P.
 PR 17-NOV-2000; 2000US-0249211P.
 PR 17-NOV-2000; 2000US-0249212P.
 PR 17-NOV-2000; 2000US-0249213P.
 PR 17-NOV-2000; 2000US-0249214P.
 PR 17-NOV-2000; 2000US-0249215P.
 PR 17-NOV-2000; 2000US-0249216P.
 PR 17-NOV-2000; 2000US-0249217P.
 PR 17-NOV-2000; 2000US-0249218P.
 PR 17-NOV-2000; 2000US-0249244P.
 PR 17-NOV-2000; 2000US-0249245P.
 PR 17-NOV-2000; 2000US-0249264P.
 PR 17-NOV-2000; 2000US-0249265P.
 PR 17-NOV-2000; 2000US-0249297P.
 PR 17-NOV-2000; 2000US-0249299P.
 PR 17-NOV-2000; 2000US-0249300P.
 PR 01-DEC-2000; 2000US-0250160P.
 PR 05-DEC-2000; 2000US-0251030P.
 PR 05-DEC-2000; 2000US-0251988P.
 PR 05-DEC-2000; 2000US-0251989P.
 PR 06-DEC-2000; 2000US-0251719P.
 PR 08-DEC-2000; 2000US-0251856P.
 PR 08-DEC-2000; 2000US-0251868P.
 PR 08-DEC-2000; 2000US-0251869P.
 PR 08-DEC-2000; 2000US-0251989P.
 PR 08-DEC-2000; 2000US-0251990P.
 PR 11-DEC-2000; 2000US-0251997P.
 PR 05-JAN-2001; 2001US-0259678P.
 XX
 PA (HUMA-) HUMAN GENOME SCI INC.

```

XX Rosen CA, Barash SC, Ruben SW,
XX WPI; 2001-483426/52.
DR N-PSDB; AAK56945.
XX
PT Nucleic acids encoding human immune/hematopoietic antigen polypeptides,
PT useful for preventing, diagnosing and/or treating cancers and metastasis.
XX
XX Claim 11; SEQ ID NO 11757; 3071pp + Sequence Listing; English.
XX
XX AAK54951 to AAK64702 encode the human immune/hematopoietic antigen (I)
XX amino acid sequences given in AAM82170 to AAM91921. (I) have cytostatic
XX activity, and can be used in gene therapy and vaccine production. (I)
XX proteins and polynucleotides may be used in the prevention, diagnosis and
XX treatment of diseases associated with inappropriate (I) expression. For
XX example, they may be used to treat disorders associated with decreased
XX expression by rectifying mutations or deletions in a patient's genome
XX that affect the activity of (I) by expressing inactive proteins or to
XX supplement the patient's own production of (I). Additionally, (I)
XX polynucleotides may be used to produce the secreted (I), by inserting the
XX nucleic acids into a host cell and culturing the cell to express the
XX protein. (I) proteins and polynucleotides may be used to prevent,
XX diagnose and treat immune/hematopoietic-related diseases, especially
XX cancers and cancer metastases of hematopoietic-derived cells. AAK64703
XX to AAK87694 represent human immune/hematopoietic antigen genomic
XX sequences from the present invention. AAK54942 to AAK54950 and AAM82169
XX represent sequences used in the exemplification of the present invention
XX
SQ Sequence 60 AA;

Query Match 38.6%; Score 61; DB 4; Length 60;
Best Local Similarity 45.0%; Pred. No. 0.54;
Matches 9; Conservative 5; Mismatches 6; Indels 0; Gaps 0;

QY 1 VFTGVLAGWGGGKFPVYPF 20
   :|:|:|:|:|:|:|:|:|
Db 34 IFTEIASGYWGSGKRFPLNF 53

RESULT 12
ADC97033
ID ADC97033 standard; protein; 373 AA.
XX
AC ADC97033;
XX
DT 01-JAN-2004 (first entry)
XX
DE E. faecium protein sequence SEQ ID 6660.
XX
XX Vaccine; urinary tract infection; bacteraemia; endocarditis; wound;
XX abdominal-pelvic infection.
XX
OS Enterococcus faecium.
XX
XX US6593275-B1.
XX
XX 24-JUN-2003.
XX
XX 30-JUN-1998; 98US-00107532.
XX
XX 02-JUL-1997; 97US-0051571P.
XX
XX 14-MAY-1998; 98US-0085598P.
XX
XX (GENO-) GENOME THERAPEUTICS CORP.
XX
XX Doucette-Stamm LA, Bush D;
XX
XX WPI; 2003-799836/75.
XX
XX N-PSDB; ADC93379.
XX
XX New isolated nucleic acid derived from Enterococcus faecium encoding an
XX Enterococcus faecium polypeptide useful for detection, prevention and

```

```

PT treatment of a pathological condition resulting from a bacterial
PT infection.
XX
XX Example 1; SEQ ID NO 6660; 243pp; English.
XX
XX The invention relates to an isolated nucleic acid derived from
XX Enterococcus faecium encoding an Enterococcus faecium polypeptide having
XX one of 10 fully defined sequences given in the (or comprising 40
XX sequential nucleotides chosen from any of the nucleic acids, its
XX complement or sequences hybridizing to it). Also included are a
XX recombinant vector comprising the nucleic acid operably linked to
XX transcription regulatory element, a cell comprising the vector and a
XX single-stranded probe comprising the nucleic acid. The nucleic acids are
XX chosen from 3654 disclosed sequences encoding 3654 disclosed proteins.
XX The nucleic acids is useful for diagnosing pathological conditions
XX resulting from E. faecium bacterial infection (e.g. urinary tract
XX infection, bacteraemia, endocarditis, wounds and abdominal-pelvic
XX infection) and for screening drugs such as agonists and antagonists. The
XX nucleic acid is useful for recombinant production of Candida albicans -
XX derived peptides or antisense polypeptides. Pharmaceutical compositions
XX and vaccines containing the nucleic acid are useful for preventing or
XX treating Enterococcus faecium infections. The present sequence represents
XX one if the disclosed E. faecium proteins.
XX
SQ Sequence 373 AA;

Query Match 34.5%; Score 54.5; DB 7; Length 373;
Best Local Similarity 47.6%; Pred. No. 28;
Matches 10; Conservative 4; Mismatches 6; Indels 1; Gaps 1;

QY 9 VMGEGGKFPVFPD-DKMSFLF 28
   :|:|:|:|:|:|:|:|:|
Db 333 IMQVGKFPVFPNKDLEFNF 353

RESULT 13
ABG20173
ID ABG20173 standard; protein; 134 AA.
XX
XX ABG20173;
XX
DT 18-FEB-2002 (first entry)
XX
DE Novel human diagnostic protein #20164.
XX
XX Human; chromosome mapping; gene mapping; gene therapy; forensic;
XX food supplement; medical imaging; diagnostic; genetic disorder.
XX
XX Homo sapiens.
XX
XX WO200175067-A2.
XX
XX 11-OCT-2001.
XX
XX 30-MAR-2001; 2001WO-US008631.
XX
XX 31-MAR-2000; 2000US-00540217.
XX
XX 23-AUG-2000; 2000US-00649167.
XX
XX (HYSE-) HYSEQ INC.
XX
XX Drmanac RT, Liu C, Tang YT;
XX
XX WPI; 2001-639362/73.
XX
XX N-PSDB; AAS84360.
XX
XX New isolated polynucleotide and encoded polypeptides, useful in
XX diagnostics, forensics, gene mapping, identification of mutations
XX responsible for genetic disorders or other traits and to assess
XX biodiversity.
XX
XX Claim 20; SEQ ID NO 50532; 103pp; English.
XX

```


DR N-PSDB; AAT78508.

XX DNA encoding Streptomyces ambofaciens platenolide synthase domain - for
PT production of spiramycin-related polyketide antibiotics.
XX

PS Claim 8; Page 66-71; 81pp; English.

CC AAW23716-W23720 represent proteins encoded by the platenolide synthase
CC gene cluster of the invention. The gene cluster is also referred to as
CC the srmg gene, and was isolated from Streptomyces ambofaciens. These
CC sequences are multi-functional proteins which direct the synthesis of the
CC polyketide platenolide. Platenolide is the basic building block of the
CC macroide antibiotic spiramycin. The DNA can be used to produce compounds
CC exhibiting antibiotic activity based on the platenolide structure,
CC including specifically the macroide antibiotic spiramycin and spiramycin
CC analogues and derivatives. Modifications of the platenolide synthase DNA
CC sequence can be made so as to change the number and type of carboxylic
CC acids incorporated into the growing polyketide chain and to change the
CC kind of post-condensation processing that is conducted

SO Sequence 1580 AA;

Query Match 33.5%; Score 53; DB 2; Length 1580;
Best Local Similarity 75.0%; Pred. No. 2e+02;
Matches 9; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 7 AGVWGGGKFFY 18
||| ||| : ||
Db 1287 AGVWGGGQSVY 1298

Search completed: September 30, 2004, 11:04:45
Job time : 78.1238 secs

This Page Blank (uspto)